Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/187345

Please be advised that this information was generated on 2019-12-04 and may be subject to change.

Anterior cervical discectomy for single level degenerative disease. Some considerations

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus Prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 20 april 2018
om 11.30 uur precies

door
Roland Dirk Donk
geboren 2 april 1956
te Venlo

CB2
C
9196

EX LIBRIS
UNIVERSITATIS
NOVIOMAGENSIS

Promotor: Prof. dr. R.H.M.A. Bartels

Copromotoren: Dr. A.J.F. Hosman
Dr. W.I.M. Verhagen (CWZ Nijmegen)

Manuscriptcommissie:

Prof. dr. M.J.R. Edwards
Prof. dr. B.J. van Royen (VU Medisch Centrum)
Prof. dr. J. Goffin (KU Leuven , België)
Dr. C.L.A.M. Vleggeert-Lankamp (Leids Universitair Medisch Centrum)
Dr. D.J. Jeszenszky (Schulthess Klinik Zürich, Zwitserland)

Content

Chapter 1	General Introduction.	7
Chapter 2	Reporting the results of meta-analyses: A plea for incorporating clinical relevance referring to an example.	13
Chapter 3	The qualification of outcome after cervical spine surgery by patients compared to the Neck Disability Index.	31
Chapter 4	What's the best surgical treatment for patients with cervical radiculopathy due to single-level degenerative disease? A randomized controlled trial.	41
Chapter 5	Factors determining a good recovery after surgery for symptomatic single level cervical degenerative disc disease.	65
Chapter 6	Symptomatic adjacent segment disease after anterior cervical discectomy for one-level degenerative disc disease.	75
Chapter 7	An assessment of the most reliable method to estimate the sagittal alignment of the cervical spine: analysis of a prospective cohort of 138 cases.	87
Chapter 8	Cervical sagittal alignment after different anterior discectomy procedures for single level cervical degenerative disc disease: randomized controlled trial.	99
Chapter 9	Summary and Considerations.	115

Chapter 1

General introduction

General Introduction

Cervical disc degeneration is frequently associated with arm and/or neck pain. Compression of the nerve root might cause signs and symptoms of radiculopathy: arm pain radiating from the neck, tingling, numbness, and muscle weakness[1].

Cervical radiculopathy due to cervical disc degeneration is frequently encountered in daily practice with an average incidence of 83 per 100,000. The age-specific annual incidence rate per 100,000 reached a peak among those aged 50-54[2].

In most instances conservative treatment is the first option with satisfactory results[3]. When conservative treatment fails, surgical treatment can be considered. Since the description by Cloward[4], Smith and Robinson[5], Hirsch[6], and Dereymaeker and Mulier[7], the anterior approach became the favored approach to surgically treating cervical degenerative disk disease.

The anterior approach is known as anterior cervical discectomy (ACD), and the major goal is decompression of the nerve root by removing the disc, disc fragments, and osteophytes. It can be performed without any additional measures such as those reported by Hirsch or with fusion. The ACD with fusion (ACDF) with autograft as described by Cloward, Smith and Robinson, and Dereymaeker and Mulier became the gold standard[8]. However, ACD is still considered a good alternative for single-level cervical degenerative disease[9].

Currently several alternatives are available for autograft from the iliac crest. A stand-alone cage has been proven to be a reasonable option[8], and has been adopted by many spine surgeons.

In the last decade of the previous century, attention was paid to the possible disadvantages of fusion. It was thought to contribute to accelerated degeneration of the disc of the adjacent segment: adjacent segment disease (ASD). The paper by Hillibrand[10], in which he calculated an annual rate for the development of ASD of 2.9%, is often cited.

Although his calculation is debatable[11], this report was used to justify the development and clinical implementation of a new implant, the disc prosthesis or arthroplasty (ACDA) in the late 1990s. In 2002 Goffin et al presented the preliminary results of the Bryan disc prosthesis[12].

At that time confusion existed about the best surgical treatment for single-level degenerative disc disease by an anterior approach: ACD, ACDF or ACDA. This was the reason to design a randomized controlled trial (RCT) comparing ACD, ACDF with stand-alone cage, and ACDA (Bryan prosthesis)[13]. The first patient for this trial was included in October 2003.

Several RCTs comparing ACDF with plate and ACDA were started at that time. In 2010 we published the first meta-analysis of the studies that had presented their results[14]. Since we found no difference we could not justify continuation of our trial especially since the costs of the implant significantly exceeded those of a cage. Therefore, the RCT was concluded prematurely. However,

we felt an obligation towards our patients to evaluate the results. These results contributed to this thesis.

In the second chapter (Chapter 2) attention is paid to the importance of reporting minimal clinically important differences while reporting the results of meta-analyses. Recent meta-analyses comparing ACDF and arthroplasty were used as an example.

The definition of a good result after an ACD procedure is important for the patient but also for the physician. Ideally this is based on a patient-reported outcome measurement (PROM). In **Chapter 3** NDI is used to define a good outcome.

The clinical results of our RCT are described in **Chapter 4**. Since we were also interested in which factors were correlated with a good outcome, we performed separate analyses. These are presented in **Chapter 5.**

ASD is still a matter of subject of debate. Due to the long follow-up and the high response rate of more than 98% we investigated the occurrence of ASD that is clinically relevant. We could calculate an annual rate for clinically relevant ASD. These results are presented in **Chapter 6**.

The influence of surgery on the sagittal angle of the level of interest but also the angle from C2 to C7 as a measure of global sagittal alignment is currently subject of debate. In order to develop a reliable method to assess the sagittal alignment of the cervical spine, we reviewed our cervical radiographs that were obtained according to the protocol of the RCT. This method is described in **Chapter 7**. The actual angles were measured and the sagittal alignment of the cervical spine assessed by two investigators. These results are described in **Chapter 8**. In the last chapter (**Chapter 9**) the results are summarized and some considerations made.

References

1. Kuijper B, Tans JT, Schimsheimer RJ, et al. Degenerative cervical radiculopathy: diagnosis and conservative treatment. A review. *Eur J Neurol.* Jan 2009;16(1):15-20.

2. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT. Epidemiology of cervical radiculopathy. A population-based study from Rochester, Minnesota, 1976 through 1990. *Brain.* Apr 1994;117 (Pt 2):325-335.

3. Kuijper B, Tans JT, Beelen A, Nollet F, de Visser M. Cervical collar or physiotherapy versus wait and see policy for recent onset cervical radiculopathy: randomised trial. *BMJ.* Oct 07 2009;339:b3883.

4. Cloward RB. The anterior approach for removal of ruptured cervical disks. *J Neurosurg.* Nov 1958;15(6):602-617.

5. Smith GW, Robinson RA. The treatment of certain cervical-spine disorders by anterior removal of the intervertebral disc and interbody fusion. *J Bone Joint Surg Am.* Jun 1958;40-A(3):607-624.

6. Hirsch C, Wickbom I, Lidstroem A, Rosengren K. Cervical-Disc Resection. A Follow-up of Myelographic and Surgical Procedure. *J Bone Joint Surg Am.* Dec 1964;46:1811-1821.

7. Dereymaeker A, Mulier J. [Vertebral fusion by a ventral approach in cervical intervertebral disk disorders]. *Rev Neurol (Paris).* Dec 1958;99(6):597-616.

8. Jacobs W, Willems PC, Kruyt M, et al. Systematic review of anterior interbody fusion techniques for single- and double-level cervical degenerative disc disease. *Spine (Phila Pa 1976).* Jun 15 2011;36(14):E950-960.

9. de Rooij JD, Gadjradj PS, Soria van Hoeve JS, Harhangi BS. Anterior cervical discectomy without fusion for a symptomatic cervical disk herniation. *Acta Neurochir (Wien).* Jul 2017;159(7):1283-1287.

10. Hilibrand AS, Carlson GD, Palumbo MA, Jones PK, Bohlman HH. Radiculopathy and myelopathy at segments adjacent to the site of a previous anterior cervical arthrodesis. *J Bone Joint Surg Am.* Apr 1999;81(4):519-528.

11. Donk RB, R.H.M.A. Adjacent disc degeneration in the cervical spine: personal data and a critical reappraisal of the literature. *The Internet Journal of Spine Surgery.* 2012;6(2):1-8.

12. Goffin J, Casey A, Kehr P, et al. Preliminary clinical experience with the Bryan Cervical Disc Prosthesis. *Neurosurgery.* Sep 2002;51(3):840-845; discussion 845-847.

13. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord.* 2006;7:85.

14. Bartels RH, Donk R, Verbeek AL. No justification for cervical disk prostheses in clinical practice: a meta-analysis of randomized controlled trials. *Neurosurgery.* Jun 2010;66(6):1153-1160; discussion 1160.

Chapter 2

Reporting the results of meta-analyses: A plea for incorporating clinical relevance referring to an example.

Ronald H.M.A. Bartels, Roland D. Donk, Wim I.M. Verhagen, Allard J.F. Hosman, André L.M. Verbeek

Published in The Spine Journal (2017), PII: S1529-9430(17)30211-5
http://dx.doi.org/doi: 10.1016/j.spinee.2017.05.019.

Abstract

Background Context

The results of meta-analyses are frequently reported, but understanding and interpreting them is difficult for both clinicians and patients. Statistical significances are presented without referring to values that imply clinical relevance.

Purpose

This study aimed to use the minimal clinically important difference (MCID) to rate the clinical relevance of a meta-analysis.

Study design

This study is a review of the literature.

Patient sample

This study is a review of meta-analyses relating to a specific topic, clinical results of cervical arthroplasty.

Outcome measure

The outcome measure used in the study was the MCID.

Methods

We performed an extensive literature search of a series of meta-analyses evaluating a similar subject as an example. We searched in PubMed and Embase through August 9, 2016 and found articles concerning meta-analyses of the clinical outcome of cervical arthroplasty compared to that of anterior cervical discectomy with fusion in cases of cervical degenerative disease. We evaluated the analyses for statistical significance and their relation to MCID. MCID was defined based on results in similar patient groups and a similar disease entity reported in the literature.

Results

We identified 21 meta-analyses, only one of which referred to MCID. However, the researchers used an inappropriate measurement scale and, therefore, an incorrect MCID. The majority of the conclusions were based on statistical results without mentioning clinical relevance.

Conclusions

The majority of the articles we reviewed drew conclusions based on statistical differences instead of clinical relevance. We recommend introducing the concept of MCID while reporting the results of a meta-analysis, as well as mentioning the explicit scale of the analyzed measurement.

List of abbreviations

ACDF	anterior cervical discectomy with fusion
MCID	minimal clinically important difference
MCS	mental component summary
NDI	neck disability index
NRS	numeric rating scale
OR	odds ratio
PCS	physical component summary
RR	risk ratio
SF 36	short form 36
SMD	standardized mean difference
VAS	visual analogue scale
WMD	weighted mean difference

Introduction

Meta-analyses combine the outcomes of multiple but similar studies to derive a pooled estimate with a statistically stronger conclusion than each study could achieve separately. A well-known complication is that the clinical relevance of the problem is forgotten while searching for statistical significance, claiming an effect that is not beneficial for the patient and may also increase the costs of health care.

The concept of minimal clinically important difference (MCID) is used to indicate whether a treatment might be beneficial to a patient. Although MCID is an attractive concept, pitfalls are manifold [1]. MCID is influenced by the choice of measurement, trial design and study population. Although Katz, Paillard and Ekman do not advise applying MCID to a group [1], it can be useful in determining whether a treatment is meaningful for a patient [2-4]. Meta-analyses of randomized controlled trials (RCTs) investigate nearly similar study populations, equal measurements and trial designs. Therefore, a chosen MCID that has been constructed within a similar group of patients with the same disease as the subject of interest can contribute to estimating the subject of the analysis.

It is very tempting to use MCID as a value to rate the clinical relevance of a meta-analysis, since a sound alternative is still lacking. Meta-analyses about cervical arthroplasty can serve as an example. In their daily practice, physicians frequently encounter cervical degenerative disc disease which causes radicular symptoms and signs. Although a conservative attitude is appropriate because patients will recover in most instances, surgery is sometimes warranted. Several surgical options exist: anterior and posterior approaches are both possible, though the anterior approach is most often taken. The gold standard is anterior cervical discectomy with fusion (ACDF). A more recent technique is arthroplasty using a cervical disc prosthesis to maintain mobility and to reduce the incidence of adjacent disc disease.

Many reports of RCTs at different follow-up periods have been published. The main outcome is the clinical effect, measured by several patient-reported outcome measurements. Meta-analyses have been published to analyze the pooled results. Every update of a RCT because of longer follow up is followed by a series of new meta-analyses. Although statistical significances are reported, the clinical relevance is emphasized less often.

The main goal of this study is to investigate whether the results of the meta-analyses comparing arthroplasty with ACDF were reported while addressing clinical relevance, and whether the conclusions were supported by the data.

Methods

Although this study is neither a systemic review nor a meta-analysis, we will follow the PRISMA statement [5] to provide a complete and clear report of the search strategy.

The literature was searched using PubMed and Embase through August 9, 2016. The following search string was used: ((disc OR disk) AND cervical) AND (arthroplasty OR prosthesis) AND meta-analysis. Two investigators (RB and RD) then read the articles' titles and abstracts. Only those meta-analyses that address cervical arthroplasty because of one- or two-level degenerative

disease were included. They also needed to address clinical outcome measurements such as the Neck Disability Index (NDI), Visual Analogue Scale (VAS) arm, VAS neck, and the Physical Component Summary (PCS) of the Short Form 36 (SF 36). Articles solely reporting radiological outcome and/or subsequent surgeries were excluded, as were those with hybrid solutions as a subject. If RB and RD disagreed, AV was consulted. Of interest for this study were the mean value, the 95% confidence interval (95% CI) and the number of included patients. We only used the data from the latest follow up within a study.

As a threshold for clinical relevance, we chose an MCID for NDI, VAS arm, VAS Neck, and SF 36 PCS based on the literature. The search for the most appropriate MCID was conducted according to Katz et al.'s [3] suggestions. The NDI is expressed as an absolute number (scale: 0–50) or as a percentage, depending on the number of questions completed (scale: 0–100%). The MCID is set at 10% or 20% [6]. Because an MCID of 7.5% or 15% is also often used, these were included in the comparison.

The VAS is measured by drawing a mark on a 100-mm line. The distance from zero is the value, and can therefore vary from 0 to 100. The MCID for cervical radicular symptoms is twofold: for neck pain it is 4.1 and for arm pain 2.6 (scale: 0–10), with an odds ratio(OR) of41 and 26 respectively (scale: 0–100) [7].

The numeric rating scale (NRS) is different from the VAS because the patient is asked to rate their own pain on a scale of 0 to 10. The MCID for the NRS for the neck is set at 4.6 (scale: 0–10) or 46 (scale: 0–100) [8]. After an extensive search, we found no articles reporting the MCID for NRS arm. Therefore, we set the level at 2.0 from a pragmatic point of view [9], taking into account the similarity with VAS.

The Short Form 36 is a 36-item questionnaire used to measure quality of health. Two scores can be extracted: the mental component summary relating to mental health and the PCS reflecting physical well-being. The MCID for SF 36 PCS is set at 8.1 [7].

If a standardized mean difference was used, the standard deviation was calculated based on the 95% confidence interval. Afterwards, the MCID was divided by the SD to standardize the MCID. If an odds ratio (OR) or relative risk (RR) were reported, we calculated the log of the MCID. In case of a weighted mean difference (WMD), we used the MCID as a discrete value. The values were represented as positive discrete values. If a weighted mean difference or standardized mean difference has a negative value, it should be smaller than the negative value of the MCID to imply clinical relevance.

Finally, we graded the conclusions: positive if they suggested superiority of arthroplasty compared to ACDF, neutral if a preference was not mentioned or negative if the value of the arthroplasty was explicitly questioned. These ratings were made from the perspective of a non-native English speaking person. Because comparisons between studies were not the objective of present study, we did not compute any reliability statistics.

Results

Of the 32 articles retrieved by the search, we selected 21 [10-30]. The flowchart diagram in accordance with the PRISMA statement is represented in Figure 1.

Two articles mentioned an NRS score for pain [15,30] but, after revising the included articles in the meta-analyses, it became clear that a VAS was meant. Therefore, an MCID in relation to NRS could not be investigated.

Only Fallah et al.'s [12] article mentioned MCID. The authors correlated the outcome to the MCID as reported in the literature. However, they did not use the MCID for the correct scale in the investigated comparisons. For example, an MCID of 7.5 was used for the NDI, corresponding to an NDI scale from 0 to 50. However, the included articles used the NDI % scale (ranging from 0–100). This increased the MCID to 15, making the calculated differences no longer clinically relevant. The same holds true for the VAS scores; in the original articles, they were measured on a scale from 0 to 100 instead of 0 to 10.

The remaining articles did not mention the MCID and only reported whether a comparison revealed statistical significance or not (see Tables 1–5 for NDI, VAS arm, VAS neck and SF 36 PCS). In those tables, the MCID are represented as reported in the literature. Comparing these values with the reported differences or ORs or relative risk, we found that almost none of them reached clinical relevance depending upon the outcome scale used. The scales in the meta-analyses did not always correspond with the scales in the original article. For example, an MCID for a VAS on a 1 to 100 scale is different than a VAS on a 1 to 10 scale. All the articles except Fallah et al. [12] formulated the conclusions based on strictly statistical results. Fourteen (66.7%) of the articles ended with positive conclusions [10,13,14,18,19,21-23,25-29], six (28.6%) drew neutral conclusions [12,15,16,20,24,30] and one (4.7%) drew a negative conclusion [11].

Discussion

Meta-analyses are frequently published, as they are a well-known and sound method of pooling the results of different studies to find the best available evidence [31]. However, they are difficult for clinicians and patients to interpret. Statistical significance does not automatically equal clinical relevance [3,32].

Relating calculated differences or ratios to a value that might imply clinical relevance would contribute to better understanding and interpretation of the results of a meta-analysis. If costly treatments are involved, it would also be helpful from a societal perspective. A value like the MCID would be welcome in evaluating whether results are meaningful for a patient.

MCID was developed to characterize the smallest amount of change that the patient perceived as meaningful [4,33]. However, it varies widely depending on various factors [1,34]. Despite its shortcomings, MCID is currently the best potential benchmark for evaluating the results of meta-analyses of RCTs, because the study populations, study designs and outcome measurements

are similar. A very important condition is that the researchers (preferably the care providers and patients) define the level of MCID.

In our example, MCID was defined based on the results in similar patient groups with the same disease entity. Therefore, we feel confident that the MCID was not substantially different. A remarkable finding was that the conclusions were based on statistical significance that did not appear to correspond with clinical relevance. Furthermore, caution is needed when applying an MCID, since it is not always clear which scale is used for the outcome measurement.

Our review of meta-analyses limited to a single pathology could be considered a limitation of this study. However, we feel confident that focusing on one subject increased the impact of the example. Recently, many meta-analyses on the subject of interest have been published with similar outcome measurements. Therefore, we could clearly show that many recent studies with a similar design, the same subject of interest and the same outcome measurements did not address clinical importance.

The lack of mentioning MCID, which we assumed to be a reflection of clinical relevance, should be a point of interest when reviewing or interpreting the results of meta-analyses that report composite or indirect outcome measurements (e.g., NDI, SF 36, VAS or NRS). The understanding or interpretation of discrete outcome measurements such as mortality or infection rate can be more straightforward without any need to provide a value for MCID. Although we assume that introducing MCID will improve the understanding and interpretation of results of meta-analyses that report on non-discrete outcome measurements, we would like to stress that we did not prove it in this study. We have only shown that it was not mentioned.

An attractive alternative is the introduction of the number to treat (NNT) as some already have suggested [35,36]. The number to treat expresses the number of patients that must be treated to achieve an event of clinical impact.

It is easy to calculate NNT. To compare the effect of a treatment in two groups, start with the frequency of the outcome of interest for each group separately. The difference between these two frequencies is the attributable risk (AR). NNT equals 1/AR.

The interpretation of NNT is very intuitive for clinicians: NNT expresses the number of patients that you need to treat to establish the outcome of interest if the patient is treated by one method instead of another. If the outcome of interest is a good outcome after treatment, the NNT is ideally low.

The presentation of frequencies to calculate AR and eventually NNT is a prerequisite. However, as in our example, meta-analyses often report weighted or standardized means. Therefore, the use of NNT is restricted to meta-analyses that mention proportions or frequencies.

To improve the understanding and interpretation of the results of meta-analyses, we recommend introducing the concept of MCID to analyses that report on non-discrete outcome measurements. The values reported in the literature for similar diseases and study populations could be used as a guide. Ideally, researchers, physicians and patients would define these values. The measurement scales of the analyses should be mentioned explicitly and separately in the presentation of meta-analyses.

Figure legends

Figure 1: PRISMA flowchart

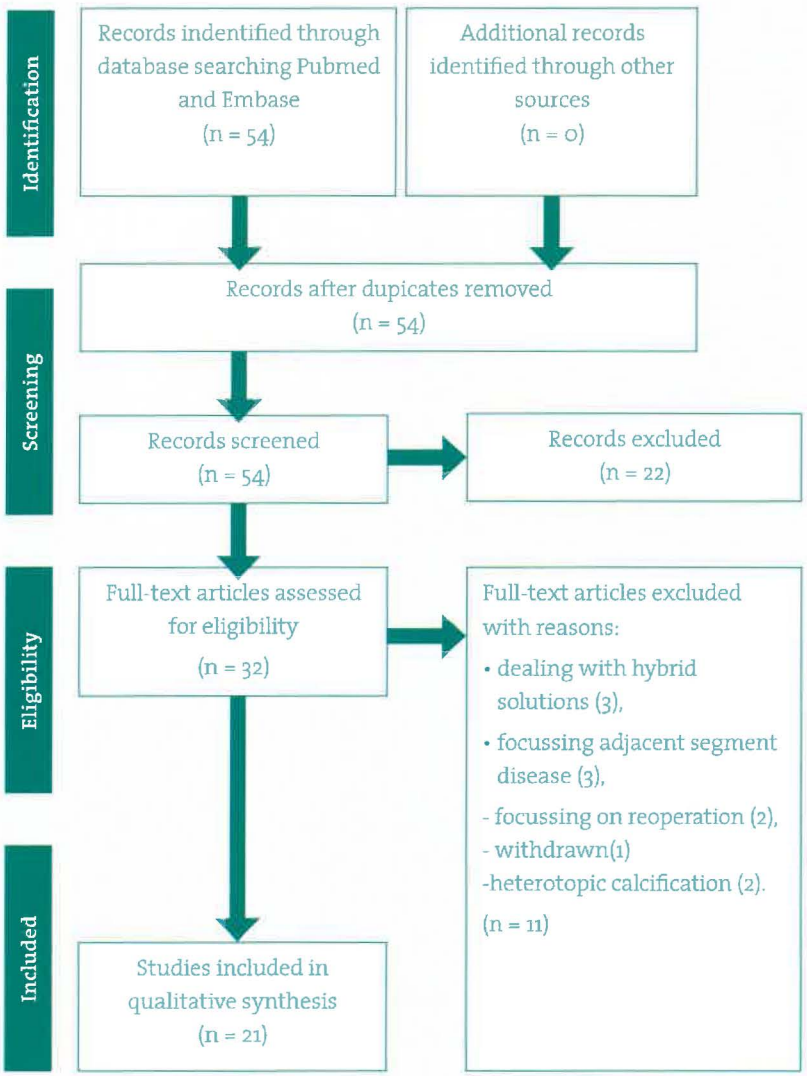


Table 1: NDI and the estimated MCID at four levels.

Author	Number of patients	SMD/ WMD/ OR1	Mean value	95% CI lowest2	95% CI highest3	SD4	MCID			
							NDI =7.5	NDI = 10	NDI = 15%	NDI =20%
Zhu et al. 2016	852	SMD	-0.23	-0.36	-0.09	1.94	3.9	5.2	7.7	10.3
Zou et al. 2016	502	SMD	0.31	0.12	0.5	2.17	3.5	4.6	6.9	9.2
Kuang et al. 2016	588	WMD	-1.53	-3.8	0.73		7.5	10	15	20
Hu et al. 2016	1713	WMD	-6.68	-9.17	-4.2		7.5	10	15	20
Yao et al. 2016	1495	OR	0.88	0.73	1.3		0.9	1.0	1.2	1.3
Aragones et al. 2015	1005	WMD	0.88	0.25	1.51		7.5	10	15	20
Wu et al. 2015	921	WMD	-6.59	-6.93	-6.26		7.5	10	15	20
Muheremu et al. 2015[5]	837	SMD	-1.11	-1.87	-0.35	11.22	0.7	0.9	1.3	1.8
Zhang et al. 2015	590	SMD	-0.24	-0.38	-0.09	1.73	4.3	5.8	8.7	11.5
Ren et al. 2014	653	WMD	5.49	2.79	8.2		7.5	10	15	20
Luo et al. 2015	907	SMD	-1	-5.28	3.28	65.76	0.1	0.2	0.2	0.3
Li et al. 2015	1601	RR	0.95	0.91	1		0.8	1.0	1.18	1.3
Gao et al. 2015	760	WMD	-3.38	-8.71	1.95		7.5	10	15	20
Yin et al. 2013	704	SMD	-0.27	-0.42	-0.11	2.03	3.7	4.9	7.4	9.9
Xing et al. 2013	1374	WMD	-3.81	-8.12	0.51		7.5	10	15	20
Fallah et al. 2012	1174	WMD	-3.03	-6.21	0.16		7.5	10	15	20
McAfee et al. 2012	1191	OR	0.786	0.589	1.05		0.9	1.0	1.2	1.3
Jiang et al. 2012	1046	SMD	0.01	-0.25	0.27	4.29	1.8	2.3	3.5	4.7
Yu et al. 2011	720	SMD	-0.09	-0.44	0.27	4.79	1.6	2.1	3.1	4.2
Bartels et al. 2010	1533	OR	0.794	0.641	0.984		0.9	1.0	1.2	1.3

MCID, minimal clinically important difference; NDI, Neck Disability Index

[1] SMD: standardized mean difference; WMD: weighted mean difference; OR: odds ratio or relative risk
[2] Lowest border of 95% Confidence Interval (CI)
[3] Highest border of 95% Confidence Interval (CI)
[4] Calculated standard deviation based on number of patients, mean and 95% CI from reviewed study
[5] Used NDI %

Table 2: NDI statistical relevance versus clinical relevance. NDI expressed as absolute scores (scale 0-50) and percentage scores (scale 0-100)

Author	Mean value	Statistical significance*	NDI = 7.5		NDI = 10		NDI=15%		NDI=20%	
			MCID	CR*	MCID	CR*	MCID	CR*	MCID	CR*
Zhu et al. 2016	-0.23	+	3.9	-	5.2	-	7.7	-	10.3	-
Zou et al. 2016	0.31	+	3.5	-	4.6	-	6.9	-	9.2	-
Kuang et al. 2016	-1.53	-	7.5	-	10	-	15	-	20	-
Hu et al. 2016	-6.68	+	7.5	-	10	-	15	-	20	-
Yao et al. 2016	0.88	-	0.9	-	1.0	-	1.2	-	1.3	-
Aragones et al. 2015	0.88	+	7.5	-	10	-	15	-	20	-
Wu et al. 2015	-6.59	+	7.5	-	10	-	15	-	20	-
Muheremu et al. 2015	-1.11	+	0.7	+	0.9	+	1.3	-	1.8	-
Zhang et al. 2015	-0.24	+	4.3	-	5.8	-	8.7	-	11.5	-
Ren et al. 2014	5.49	+	7.5	-	10	-	15	-	20	-
Luo et al. 2015	-1	-	0.1	-	0.2	-	0.2	-	0.3	-
Li et al. 2015[5]	0.95	-	0.8	+	1.0	-	1.18	-	1.3	-
Gao et al. 2015	-3.38	-	7.5	-	10	-	15	-	20	-
Yin et al. 2013	-0.27	+	3.7	-	4.9	-	7.4	-	9.9	-
Xing et al. 2013	-3.81	-	7.5	-	10	-	15	-	20	-
Fallah et al. 2012	-3.03	-	7.5	-	10	-	15	-	20	-
McAfee et al. 2012	0.786	-	0.9	-	1.0	-	1.2	-	1.3	-
Jiang et al. 2012	0.01	-	1.8	-	2.3	-	3.5	-	4.7	-
Yu et al. 2011	-0.09	-	1.6	-	2.1	-	3.1	-	4.2	-
Bartels et al. 2010	0.794	+	0.9	-	1.0	-	1.2	-	1.3	-

CR, clinical relevance; NDI, Neck Disability Index; MCID, minimal clinically important difference.

** If the mean value exceeds the value of MCID as calculated in Table 1, the result is considered clinical valuable (+). Otherwise, it is not (-).*

Table 3: Visual Analogue Scale arm.

Author	Number of patients	SMD/ WMD/ OR1	Mean value	95% CI lowest2	95% CI highest3	SD4	MCID 26	MCID 2.6	Statistical significance	MCID 26 + MCID 2.6 combined	Clinical relevance5
Kuang et al. 2016	437	WMD	-0.23	-0.61	0.16		26	2.6	-	MCID 2.6	-
Hu et al. 2016	710	WMD	-3.72	-7.48	0.04		26	2.6	-	MCID 26	-
Aragones et al. 2015	783	WMD	0.77	0.59	0.96		26	2.6	+	MCID 2.6	-
Wu et al. 2015	714	WMD	-4.92	-7.9	-1.94		26	2.6	+	MCID 26	-
Rao et al. 2015	637	WMD	0.04	-0.23	0.31		26	2.6	-	Combined	MCID 26 – MCID 2.6 +
Muheremu et al. 2015	557	SMD	-1.03	-1.74	-0.31	8.55	3.0	0.3	+	MCID 2.6	-
Ren et al. 2014	577	WMD	9.19	6.57	11.81		26	2.6	+	MCID 26	-
Luo et al. 2015	484	SMD	-3.23	-6.48	0.02	36.48	0.7	0.1	-	MCID 26	-
Li et al. 2015	87	WMD	0.18	0.04	0.33		26	2.6	+	MCID 2.6	-
Gao et al. 2015	793	WMD	-1.46	-2.48	-0.44		26	2.6	+	Combined	-
Yin et al. 2013	588	SMD	-0.79	-1.84	0.25	12.99	2.0	0.2	-	MCID 26	-
Xing et al. 2013	395	WMD	-4.86	-6.42	-3.3		26	2.6	+	MCID 26	-
Fallah et al. 2012	1188	WMD	-2.88	-5.12	-0.63		26	2.6	+	MCID 26	-
Jiang et al. 2012[6]	844	SMD	0.17	-0.36	0.7	7.86	3.3	0.3	-	MCID 26	-
Jiang et al. 2012[6]	40	SMD	-2.69	-3.57	-1.81	2.8	9.2	0.9	+	MCID2.6	+
Yu et al. 2011	86	SMD	-0.21	-0.63	0.22	1.99	13.1	1.	-	MCID 26	-
Bartels et al. 2010	1114	OR	1.078	0.791	1.469		1.35	0.3	-	MCID 26	-

[1] *SMD: standardized mean difference; WMD: weighted mean difference; OR: odds ratio or relative risk*

[2] *Lowest border of 95% Confidence Interval (CI)*

[3] *Highest border of 95% Confidence Interval (CI)*

[4] *Calculated standard deviation based on number of patients, mean and 95%CI from reviewed study*

[5] *When combined values were analyzed, rating was done separately if the result differed between the two possibilities.*

[6] *Used two different scales for VAS*

Table 4: Visual Analogue Scale neck.

Authors	Number of patients	SMD/ WMD/ OR1	Mean value	95% CI lowest2	95% CI highest3	SD4	MCID 41	MCID 4.1	Statistical significance	MCID 41 + MCID 4.1 combined	Clinical relevance
Kuang et al. 2016	397	WMD	-0.19	-0.71	0.33		41	4.1	-	MCID 4.1	-
Hu et al. 2016	710	WMD	-7.61	-11.43	-3.79		41	4.1	+	MCID 41	-
Aragones et al. 2015	922	WMD	0.65	0.46	0.84		41	4.1	+	MCID 4.1	-
Wu et al. 2015	714	WMD	-8.91	-12.06	-5.77		41	4.1	+	MCID 41	-
Rao et al. 2015	647	WMD	-0.25	-0.56	0.06		41	4.1	-	Combined	-
Muheremu et al. 2015	876	SMD	-1.47	-2.15	-0.78	10.27	4.5	0.5	+	MCID 4.1 MCID 41	- +
Ren et al. 2014	577	WMD	5.42	0.21	10.63		41	4.1	+	MCID 41	-
Luo et al. 2015	484	SMD	-5.99	-10.54	-1.45		41	4.1	+	MCID 41	-
Li et al. 2015	87	WMD	-0.18	-0.01	-0.34		41	4.1	+	MCID 4.1	-
Gao et al. 2015	793	WMD	-1.99	-3.1	-0.87		41	4.1	+	Combined	-
Yin et al. 2013	588	SMD	-0.12	-0.25	0.01	1.61	28.6	2.9	-	MCID 41	-
Xing et al. 2013	395	WMD	-8.16	-10.46	-5.87		41	4.1	+	MCID 41	-
Fallah et al. 2012	1168	WMD	-6.56	-9.9	-3.22		41	4.1	+	MCID 41	-
Jiang et al. 2012	40	SMD	-1.95	-2.71	-1.18	2.45	18.8	1.9	+	MCID 41	-
Jiang et al. 2012	844	SMD	-0.12	-0.37	0.13	3.71	12.4	1.2	-	MCID4.1	-
Yu et al. 2011	86	SMD	-0.48	-0.91	-0.05	2.03	22.6	2.3	+	MCID 41	-
Bartels et al. 2010	1114	OR	0.889	0.599	1.319		1.7	0.7	+	MCID 41	-

[1] SMD: standardized mean difference; WMD: weighted mean difference; OR: odds ratio or relative risk

[2] Lowest border of 95% Confidence Interval (CI)

[3] Highest border of 95% Confidence Interval (CI)

[4] Calculated standard deviation based on number of patients, mean and 95%CI from reviewed study

[5] When combined values were analyzed, rating was done separately if the result differed between the two possibilities.

[6] Used two different scales for VAS

Table 5: Physical Component Summary of the Short Form 36

Authors	Number of patients	SMD/ WMD/ OR1	Mean value	95% CI lowest2	95% CI highest3	SD4	MCID 8.1	Statistical significance	Clinical relevance
Hu et al. 2016	707	WMD	2.67	0.94	4.4		8.1	+	-
Aragones et al. 2015	792	WMD	1.41	0.89	1.93		8.1	+	-
Wu et al. 2015	714	WMD	3.16	1.87	4.44		8.1	+	-
Muheremu et al. 2015	628	SMD	0.13	-0.36	0.61	6.27	0.8	-	-
Ren et al. 2014	590	WMD	1.91	0.94	2.89		8.1	+	-
Yin et al. 2013	587	SMD	0.07	-0.06	0.2	1.61	3.2	-	-
Fallah et al. 2012	590	WMD	-2.28	-4.17	-0.4		8.1	+	-
Bartels et al. 2010	1070	OR	1.126	0.906	1.401		0.91	-	-

[1] SMD: standardized mean difference; WMD: weighted mean difference; OR: odds ratio or relative risk

[2] Lowest border of 95% Confidence Interval (CI)

[3] Highest border of 95% Confidence Interval (CI)

[4] Calculated standard deviation based on number of patients, mean and 95%CI from reviewed study

References

1. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. *J Orthop Surg Res.* 2015;10:24.

2. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J.* Sep-Oct 2007;7(5):541-546.

3. van Tulder M, Malmivaara A, Hayden J, Koes B. Statistical significance versus clinical importance: trials on exercise therapy for chronic low back pain as example. *Spine (Phila Pa 1976).* Jul 15 2007;32(16):1785-1790.

4. Man-Son-Hing M, Laupacis A, O'Rourke K, et al. Determination of the clinical importance of study results. *J Gen Intern Med.* Jun 2002;17(6):469-476.

5. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* Jul 21 2009;6(7):e1000097.

6. Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. *Am J Phys Med Rehabil.* Oct 2010;89(10):831-839.

7. Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. *J Neurosurg Spine.* Feb 2013;18(2):154-160.

8. Pool JJ, Ostelo RW, Hoving JL, Bouter LM, de Vet HC. Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. *Spine (Phila Pa 1976).* Dec 15 2007;32(26):3047-3051.

9. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain.* Aug 2004;8(4):283-291.

10. Aragones M, Hevia E, Barrios C. Polyurethane on titanium unconstrained disc arthroplasty versus anterior discectomy and fusion for the treatment of cervical disc disease: a review of level I-II randomized clinical trials including clinical outcomes. *Eur Spine J.* Dec 2015;24(12):2735-2745.

11. Bartels RH, Donk R, Verbeek AL. No justification for cervical disk prostheses in clinical practice: a meta-analysis of randomized controlled trials. *Neurosurgery.* Jun 2010;66(6):1153-1160; discussion 1160.

12. Fallah A, Akl EA, Ebrahim S, et al. Anterior cervical discectomy with arthroplasty versus arthrodesis for single-level cervical spondylosis: a systematic review and meta-analysis. *PLoS One.* 2012;7(8):e43407.

13. Gao F, Mao T, Sun W, et al. An Updated Meta-Analysis Comparing Artificial Cervical Disc Arthroplasty (CDA) Versus Anterior Cervical Discectomy and Fusion (ACDF) for the Treatment of Cervical Degenerative Disc Disease (CDDD). *Spine (Phila Pa 1976).* Dec 2015;40(23):1816-1823.

14. Hu Y, Lv G, Ren S, Johansen D. Mid- to Long-Term Outcomes of Cervical Disc Arthroplasty versus Anterior Cervical Discectomy and Fusion for Treatment of Symptomatic Cervical Disc Disease: A Systematic Review and Meta-Analysis of Eight Prospective Randomized Controlled Trials. *PLoS One.* 2016;11(2):e0149312.

15. Jiang H, Zhu Z, Qiu Y, Qian B, Qiu X, Ji M. Cervical disc arthroplasty versus fusion for single-level symptomatic cervical disc disease: a meta-analysis of randomized controlled trials. *Arch Orthop Trauma Surg.* Feb 2012;132(2):141-151.

16. Kuang L, Chen Y, Wang B, Li L, Lu G. Cervical Disk Arthroplasty Versus Anterior Cervical Decompression and Fusion for the Treatment of 2-Level Cervical Spondylopathy: A Systematic Review and Meta-analysis. *Clin Spine Surg.* Jun 10 2016.

17. Li GL, Hu JZ, Lu HB, Qu J, Guo LY, Zai FL. Anterior cervical discectomy with arthroplasty versus anterior cervical discectomy and fusion for cervical spondylosis. *J Clin Neurosci.* Mar 2015;22(3):460-467.

18. Luo J, Huang S, Gong M, et al. Comparison of artificial cervical arthroplasty versus anterior cervical discectomy and fusion for one-level cervical degenerative disc disease: a meta-analysis of randomized controlled trials. *Eur J Orthop Surg Traumatol.* Jul 2015;25 Suppl 1:S115-125.

19. McAfee PC, Reah C, Gilder K, Eisermann L, Cunningham B. A meta-analysis of comparative outcomes following cervical arthroplasty or anterior cervical fusion: results from 4 prospective multicenter randomized clinical trials and up to 1226 patients. *Spine (Phila Pa 1976).* May 15 2012;37(11):943-952.

20. Muheremu A, Niu X, Wu Z, Muhanmode Y, Tian W. Comparison of the short- and long-term treatment effect of cervical disk replacement and anterior cervical disk fusion: a meta-analysis. *Eur J Orthop Surg Traumatol.* Jul 2015;25 Suppl 1:S87-100.

21. Rao MJ, Nie SP, Xiao BW, Zhang GH, Gan XR, Cao SS. Cervical disc arthroplasty versus anterior cervical discectomy and fusion for treatment of symptomatic cervical disc disease: a meta-analysis of randomized controlled trials. *Arch Orthop Trauma Surg.* Jan 2015;135(1):19-28.

22. Ren C, Song Y, Xue Y, Yang X. Mid- to long-term outcomes after cervical disc arthroplasty compared with anterior discectomy and fusion: a systematic review and meta-analysis of randomized controlled trials. *Eur Spine J.* May 2014;23(5):1115-1123.

23. Wu AM, Xu H, Mullinix KP, et al. Minimum 4-year outcomes of cervical total disc arthroplasty versus fusion: a meta-analysis based on prospective randomized controlled trials. *Medicine (Baltimore).* Apr 2015;94(15):e665.

24. Xing D, Ma XL, Ma JX, Wang J, Ma T, Chen Y. A meta-analysis of cervical arthroplasty compared to anterior cervical discectomy and fusion for single-level cervical disc disease. *J Clin Neurosci.* Jul 2013;20(7):970-978.

25. Yao Q, Liang F, Xia Y, Jia C. A meta-analysis comparing total disc arthroplasty with anterior cervical discectomy and fusion for the treatment of cervical degenerative diseases. *Arch Orthop Trauma Surg.* Mar 2016;136(3):297-304.

26. Yu L, Song Y, Yang X, Lv C. Systematic review and meta-analysis of randomized controlled trials: comparison of total disk replacement with anterior cervical decompression and fusion. *Orthopedics.* Oct 2011;34(10):e651-658.

27. Zhang Y, Liang C, Tao Y, et al. Cervical total disc replacement is superior to anterior cervical decompression and fusion: a meta-analysis of prospective randomized controlled trials. *PLoS One.* 2015;10(3):e0117826.

28. Zhu Y, Tian Z, Zhu B, Zhang W, Li Y, Zhu Q. Bryan Cervical Disc Arthroplasty Versus Anterior Cervical Discectomy and Fusion for Treatment of Cervical Disc Diseases: A Meta-analysis of Prospective, Randomized Controlled Trials. *Spine (Phila Pa 1976).* Jun 2016;41(12):E733-741.

29. Zou S, Gao J, Xu B, Lu X, Han Y, Meng H. Anterior cervical discectomy and fusion (ACDF) versus cervical disc arthroplasty (CDA) for two contiguous levels cervical disc degenerative disease: a meta-analysis of randomized controlled trials. *Eur Spine J.* Jun 17 2016.

30. Yin S, Yu X, Zhou S, Yin Z, Qiu Y. Is cervical disc arthroplasty superior to fusion for treatment of symptomatic cervical disc disease? A meta-analysis. *Clin Orthop Relat Res.* Jun 2013;471(6):1904-1919.

31. Parmar MK, Stewart LA, Altman DG. Meta-analyses of randomised trials: when the whole is more than just the sum of the parts. *Br J Cancer.* Aug 1996;74(4):496-501.

32. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: Clinical versus statistical significance. *Perspect Clin Res.* Jul-Sep 2015;6(3):169-170.

33. Michener LA, Snyder Valier AR, McClure PW. Defining substantial clinical benefit for patient-rated outcome tools for shoulder impingement syndrome. *Arch Phys Med Rehabil.* Apr 2013;94(4):725-730.

34. Theodore BR. Methodological problems associated with the present conceptualization of the minimum clinically important difference and substantial clinical benefit. *Spine J.* Jun 2010;10(6):507-509.

35. Citrome L, Ketter TA. When does a difference make a difference? Interpretation of number needed to treat, number needed to harm, and likelihood to be helped or harmed. *Int J Clin Pract.* May 2013;67(5):407-411.

36. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ.* Feb 18 1995;310(6977):452-454.

Chapter 3

The qualification of outcome after cervical spine surgery by patients compared to the Neck Disability Index.

Roland Donk, Andre Verbeek, Wim Verhagen, Hans Groenewoud,
Allard Hosman, Ronald Bartels

Published online PLoS One. 2016 Aug 23; 11(8): e0161593 . doi: 10.1371/journal.pone.0161593

Abstract

Objective

The Neck Disability Index (NDI) is a patient self-assessed outcome measurement tool to assess disability, and that is frequently used to evaluate the effects of the treatment of neck-related problems. In individualized medicine it is mandatory that patients can interpret data in order to choose a treatment. A change of NDI or an absolute NDI is generally meaningless to a patient. Therefore, a correlation between the qualification of the clinical situation rated by the patient, and the NDI score was evaluated.

Methods

Patients who completed a NDI after anterior surgery because of symptomatic single level degenerative cervical disc disease were asked one month after completion of the NDI to qualify their clinical situation of a 5-item Likert scale varying form excellent to bad. Since a clear distinction between the categories was not possible based on the total NDI score, a ROC-curve was build, and the AUC computed in order to estimate best dichotomization in qualification of the clinical situation. The best corresponding cut-off point for the NDI total score was found by studying sensitivity and specificity for all possible cut-off points.

Results

102 patients were included. The highest AUC was obtained by dichotomizing the qualification into a group with good outcome and less-good outcome. The highest sensitivity and specificity for the dichotomized qualification as good outcome corresponded to a NDI ffi 7. Sensitivity was 81.08% and specificity was 78.57%.

Conclusions

This is the first study that correlated the qualification of the situation by the patients themselves and NDI. A NDI ffi 7 corresponded to a good outcome according to the patients. This is valuable information to inform patients in their decision for any treatment.

Introduction

The Neck Disability Index (NDI) is a frequently used, well known, and in multiple languages validated outcome measurement tool to assess self-rated disability in patients with neck pain. It can be categorized as a patient reported outcome measurement tool (PROM). The NDI is frequently used in clinical practice, but also for research purposes[1-3]. The main purpose is the quantification of the difference in pre- and post-treatment condition according the patients suffering from disabling neck pathology. The NDI addresses pain and functional items related to neck problems. It has been validated in both neck pain and, especially, whiplash patients [2,4].

Informing the patient is crucial before installing any treatment. In modern times information can be gained through many resources, but the treating physician is still very important. It has been shown that fulfillment of preoperative expectations is related to the highest post-operative satisfaction. A mismatch of disease understanding and expectation between treating physician and patient might result is a less than favorable outcome according to the patient[5-7].

The information provided by PROMs as the NDI obtained from studies can contribute in sketching expectations while informing the patient before any treatment. The most useful tools in the process of gaining information or providing it are clear clinical outcomes: mortality, infection rate etc. However, PROMs including the NDI are not reporting on a clearly defined outcome but on a combination of surrogate outcomes.

An adequate interpretation of a PROM is difficult, especially since it has been demonstrated that the language used in the questionnaires is very difficult to understand for patients. As El-Daly,I. et al. stated in their conclusion: " the majority of PROMs analyzed are written at a level that is incomprehensible to the average UK adult"[8]. The usefulness of the results of PROMs with low readability is debatable.

However, the NDI was also incorporated in the earlier mentioned study. For a correct understanding of the NDI an education level of 13-15 year-old subject was needed indicating a readability level of standard English[8]. Since the translation of the NDI into Dutch has been validated [9] we feel confident that most of the patients did understand the questions and completed their questionnaires without difficulty.

Although it seemed related, the readability of a PROM is different than interpreting the result. For example and specifically for the NDI, what information is provided to the patient if he reads or hears that a mean total score of the NDI of 8 is achieved in a group of 100 patients after a certain treatment. Information should be presented in a way that is acceptable and useful for a patient[10]. In a survey among patients with scoliosis and their careers, it was advised that the information should be user friendly and in plain language[11]. For NDI grades of disability have been defined, although these also differ and are based on clinical information and not the patients qualification[2]. A grade of disability like "none to mild disability" is not very illustrative to a patient.

Therefore, we would like to correlate the total score of the NDI with a qualitative rating by the patient themselves in a way that everyone can understand. This will contribute to understanding and decision making for patients in the future.

Methods

The STROBE statement was followed[12]. The ethical board CMO Arnhem-Nijmegen approved the study. The study has been carried out in accordance with the World Medical Association Declaration of Helsinki[13].

Patients who participated in the Procon trial Current (Controlled Trials ISRCTN41681847) [14], a comparison of different anterior cervical surgery techniques for symptomatic single level degenerative disc herniation without spinal cord involvement, and who completed a NDI were included. 142 patients participated of whom 140 completed and returned the NDI. One patient died unrelated to the trial, the other refused to return the NDI questionnaire. So, 140 patients were eligible. The mean time after surgery was 9.1 ± 1.9 years (5.6–12.2 years).

Within two months after completion of the NDI, a questionnaire was sent to the patients about the qualification of their situation regarding the neck and its related problems at that moment. Although little is known about the bias introduced by sending reminders [15], we did not send reminders or contacted the non-responders.

A five-item Likert scale was used. We did not predefine the criteria, since we were interested in the qualitative judgment of the patients themselves without any bias introduced by the researcher. The possible qualifications of their situation were: excellent, very good, good, moderate, and bad.

For statistical analyses SAS version 9.2 (SAS Institute Inc. Cary NC, USA) was used. Continuous variables are depicted as value ± standard deviation (minimum-maximum). For data analysis the Student-t test was used. Dichotomization of the patient qualifications seemed to be appropriate. To estimate which qualifications could be best combined for each possible dichotomized set, a ROC curve was build and the area under the curve (AUC) was calculated. The combination with the highest AUC was chosen. To estimate the value of NDI that corresponded best with the dichotomized outcome, the cut-off value of the total NDI with the highest sensitivity and specificity was chosen. A P value < 0.05 was assumed to be statistically significant.

Results

Of the 140 eligible patients, 102 consecutive patients completed the questionnaires (response rate: 72.9%). Mean NDI was 7.5 ± 8.6 (0-34) for the responders and 6.7 ± 8.3 for non-responders. The difference in NDI did not reach statistical significance (P=0.6). Ten patients rated their situation excellent, 33 very good, 32 good, 23 moderate, and 5 qualified their situation as bad. 73.5 % of the patients rated their situation as good or better. In Figure 1 NDI is represented in relation to the Likert qualification. It was not possible to distinct the qualifications clearly based on a total NDI score. Therefore, we decided to dichotomize qualification by the patient.

The biggest AUC was obtained by dichotomizing the qualifications in the group excellent, very good and good versus the combination of moderate and bad (AUC = 0.874). The first group

consisted of those patients with a good outcome; the patients belonging to the latter will be regarded as having a less-good outcome.

Then a ROC was constructed (Fig 2). The highest sensitivity and highest specificity for a good outcome is obtained when NDI is seven or less: sensitivity was 81.08% and specificity was 78.57%. The distribution of patients after dichotomization in relation to the NDI is shown in Table 1.

Discussion

Currently, information about any treatment is very easy accessible to patients. However, interpretation of the data is very difficult or even impossible for most patients due to lack of adequate knowledge. Surrogate outcomes are provided that are valuable for scientific purposes, but are not easily transposed to lay terms.

The NDI is a questionnaire assessed by the patient self. The NID consists of ten questions, and for each question six answers are possible. The answers are ordered starting from no disability to maximal disability. The answers are graded from zero to six, and therefore the total NDI score can vary between zero and fifty. The best outcome will be a total score of zero.

The NDI has not been uniformly divided in grades of disability[16-18]. A major concern is furthermore that the investigators predefined the qualifications of each grade. They correlated it to existing questionnaires or findings at physical examinations.

From an investigators point of view a total score of the NDI of zero would correspond with an excellent outcome. We have shown that only a proportion of the patients that rated their situation as excellent had a total NDI score of zero, whereas some patients that rated their situation as good or very good, also had a total NDI score of zero. Other (probably psychological) factors that are not taken into account in the NDI, might explain this.

Transforming a total NDI score into an expression that can easily be understood by patients will help them in making a decision about their eventual treatment, and is a contribution to individualized medicine. This is achieved not only by calculating a cut off value for the total NDI score (NDI ffi 7 versus NDI > 7), but also by dichotomizing the patients' qualification in good and less good.

Not actively motivating patients to respond might be considered a flaw of the study. However, comparison of the NDI between the group of responders and non-responders convinced us that the sample is representative. Especially when the response rate of more than 70%, that can be considered as good[19], is taking into account.

Another limitation of the study could be the lack of a pre-inquiry definition of the qualifications as rated by the patients. Therefore, the distribution of the NDI for any qualification is much wider than when the qualifications were defined prior to asking the patients. However, this would

have been again the interpretation of the researcher, whereas at this moment we are convinced that the qualifications really represented the perspective of the patient.

Finally, determining the cut off value of the NDI to consider a good or less good outcome can be subject of debate. We have chosen for a conservative approach by requesting the highest sensitivity in combination with the highest specificity. Increasing the NDI score would increase sensitivity and decrease specificity, and decreasing the NDI would induce a reverse effect creating, in our opinion, a less reliable definition of good and less outcome.

Although we did not investigated whether the patients have a better understanding of the expression of a good outcome compared to mild disability, we are convinced that the first is more appealing. From a patients perspective a total NDI score or a difference in NDI score, that is however important for scientific evaluation, is meaningless. It will not help him/her in decision-making about any treatment for neck-related problems.

In conclusion, to help the patient in the decision-making for any treatment of neck-related pathology it seems obvious that expressions should be used that are understandable. Therefore, we propose that a NDI of seven or less is qualified as a good outcome.

Fig 1: Distribution of total ND score in relation to patients' qualification

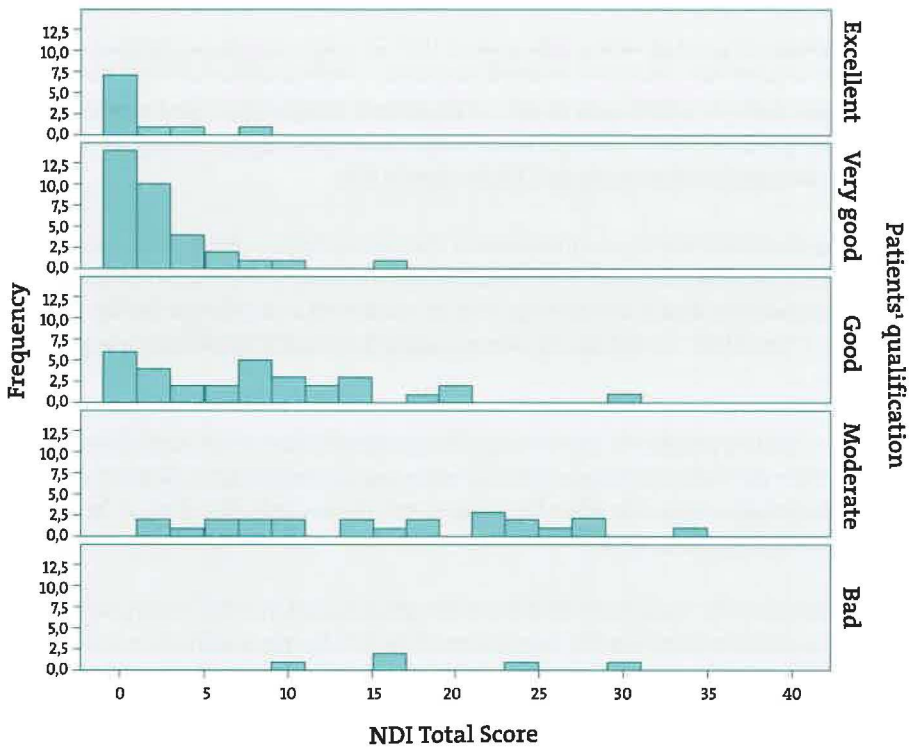


Fig 2: Figure depicting the cut-off value of the total NDI with the highest sensitivity and specificity

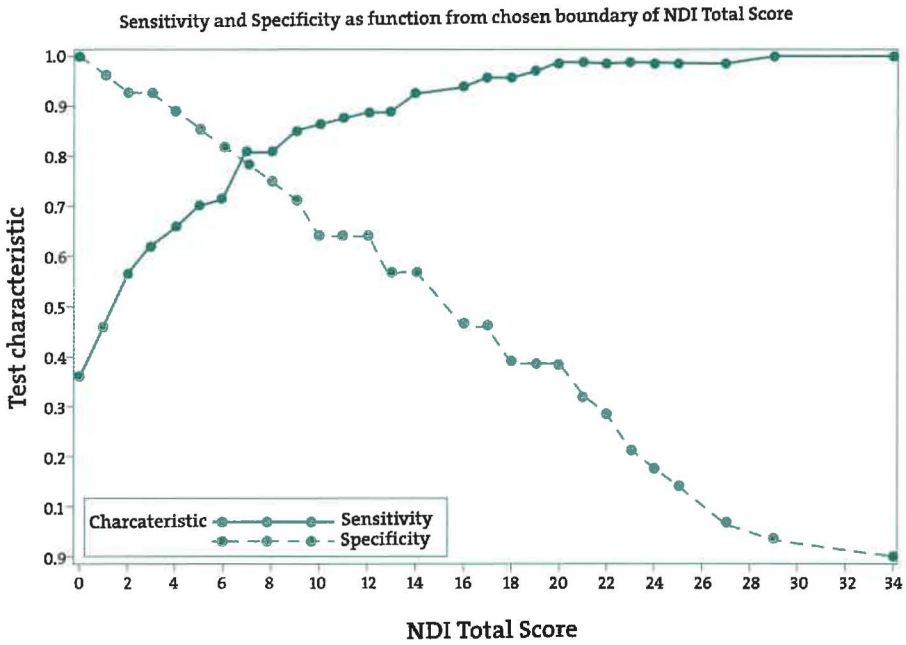


Table 1: Distribution of patients based on outcome defined as good or less-good in relation tot NDI.

	Good	Less-Good	Total
NDI ffi 7	60	14	74
NDI > 7	6	22	28
Total	66	46	102

References

1. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring the quality and effectiveness of cervical spine surgery in registry efforts: determining the most valid and responsive instruments. *Spine J.* Jun 1 2015;15(6):1203-1209.

2. MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther.* May 2009;39(5):400-417.

3. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther.* Sep 2008;31(7):491-502.

4. Howell ER. The association between neck pain, the Neck Disability Index and cervical ranges of motion: a narrative review. *J Can Chiropr Assoc.* Sep 2011;55(3):211-221.

5. Mannion AF, Junge A, Elfering A, Dvorak J, Porchet F, Grob D. Great expectations: really the novel predictor of outcome after spinal surgery? *Spine (Phila Pa 1976).* Jul 1 2009;34(15):1590-1599.

6. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 2: patient expectations and satisfaction. *Spine (Phila Pa 1976).* Jul 1 2002;27(13):1471-1476; discussion 1476-1477.

7. Soroceanu A, Ching A, Abdu W, McGuire K. Relationship between preoperative expectations, satisfaction, and functional outcomes in patients undergoing lumbar and cervical spine surgery: a multicenter study. *Spine (Phila Pa 1976).* Jan 15 2012;37(2):E103-108.

8. El-Daly I, Ibraheim H, Rajakulendran K, Culpan P, Bates P. Are patient-reported outcome measures in orthopaedics easily read by patients? *Clin Orthop Relat Res.* Jan 2016;474(1):246-255.

9. Jorritsma W, de Vries GE, Dijkstra PU, Geertzen JH, Reneman MF. Neck Pain and Disability Scale and Neck Disability Index: validity of Dutch language versions. *Eur Spine J.* Jan 2012;21(1):93-100.

10. Pellise F, Sell P, EuroSpine Patient Line Task F. Patient information and education with modern media: the Spine Society of Europe Patient Line. *Eur Spine J.* Aug 2009;18 Suppl 3:395-401.

11. Wellburn S, Bettany-Saltikov J, van Schaik P. An evaluation of web sites recommended by UK NHS consultants to patients with adolescent idiopathic scoliosis at the first point of diagnosis. *Spine (Phila Pa 1976).* Aug 15 2013;38(18):1590-1594.

12. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* Oct 20 2007;335(7624):806-808.

13. Fuson RL, Sherman M, Van Vleet J, Wendt T. The conduct of orthopaedic clinical trials. *J Bone Joint Surg Am.* Jul 1997;79(7):1089-1098.

14. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord.* 2006;7:85.

15. Tam CC, Higgins CD, Rodrigues LC. Effect of reminders on mitigating participation bias in a case-control study. *BMC Med Res Methodol.* 2011;11:33.

16. Miettinen T, Leino E, Airaksinen O, Lindgren KA. The possibility to use simple validated questionnaires to predict long-term health problems after whiplash injury. *Spine (Phila Pa 1976).* Feb 1 2004;29(3):E47-51.

17. Vernon HT. Assessment of Self-Rated Disability, Impairment, and Sincerity of Effort in Whiplash-Associated Disorder. *J Muscskel Pain.* 2000;8:155-167.

18. Sterling M, Jull G, Vicenzino B, Kenardy J. Characterization of acute whiplash-associated disorders. *Spine (Phila Pa 1976).* Jan 15 2004;29(2):182-188.

19. Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *Am J Pharm Educ.* Apr 15 2008;72(2):43.

Chapter 4

What's the best surgical treatment for patients with cervical radiculopathy due to single-level degenerative disease? A randomized controlled trial

Roland Donk, André Verbeek, Wim Verhagen, Hans Groenewoud,
Allard Hosman, Ronald Bartels

PLoS One. 2017; 12(8): e0183603. Published online 2017 Aug 29. doi: 10.1371/journal.pone.0183603

Abstract

Background

To investigate the efficacy of adding supplemental fusion or arthroplasty after cervical anterior discectomy for symptomatic mono-level cervical degenerative disease (radiculopathy), which has not been substantiated in controlled trials until now.

Methods

A randomized controlled trial is reported with 9 years follow up comparing anterior cervical anterior discectomy without fusion, with fusion by cage standalone, or with disc prosthesis. Patients suffering from symptomatic cervical disk degeneration at one level referred to spinal sections of department of neurosurgery or orthopedic surgery of a large general hospital with educational facilities were eligible. Neck Disability Index (NDI), McGill Pain Questionnaire Dutch language version (MPQ-DLV), physical-component summary (PCS), and mental-component summary (MCS) of the 36-item Short-Form Health Survey (SF-36), and re operation rate were evaluated.

Findings

142 patients between 18 and 55 years were allocated. The median follow-up was 8.9±1.9 years (5.6 to 12.2 years). The response rate at last follow-up was 98.5%. NDI at the last follow-up did not differ between the three treatment groups, nor did the secondary outcomes as MPQ-DLV and PCS or MCS from SF-36. The major improvement occurred within the first 6 weeks after surgery. Afterward, it remained stable. Eleven patients underwent surgery for recurrent symptoms and signs due to nerve root compression at the index or adjacent level.

Conclusions

This randomized trial could not detect a difference between three surgical modalities for treating a single-level degenerative disk disease. Anterior cervical discectomy without implant seems to be similar to anterior cervical discectomy with fusion by cage stand-alone or with disk prosthesis. Due to the small study sample size, this statement should be considered as inconclusive so far

Introduction

Symptomatic degeneration of a cervical intervertebral disk is encountered frequently in daily practice with irradiating pain in the arm with or with loss of sensibility or motor function as clinical presentation. The incidence varies between 0.83 and 1.79 per 1000 person-years[1,2]. In most instances, the disk will recover spontaneously without surgical intervention[3]. In case of severe pain or pain not responding to conservative treatment, surgery is a valid and effective option[4].

The anterior approach is the most often used of the surgical options. In the 1950s and 1960s, cervical anterior discectomy without (ACD) and with fusion (ACDF) were developed and propagated. Although sound evidence is still lacking for the superiority of ACDF, it serves as gold standard. Despite the high rate of recovery of non-operative therapy, an almost eight-fold increase in utilization of ACDF from 1990 to 2004 was recently reported.[5] Currently, plate fixation is considered standard for ACDF. Other fusion methods are the use of only a bone graft or a cage stand-alone. In the past two decades, another implant gained popularity, the disk prosthesis. In literature arthroplasty (ACDA) is now compared with ACDF by plate fixation. However, there has never been any definitive conclusion to the discussion of the superiority of ACDF; therefore, it is of utmost importance to complete this discussion, since more complications due to hardware failure may be involved, and the costs are significantly higher. ACDA has not been compared with ACDF with cage stand-alone, whereas the dissection for the latter is nearly the same as for ACDA, which may cause similar perioperative complication rates.

Despite statistically significant superiority, the clinical outcome after ACDA and ACDF with plate fixation is similar regarding clinical relevance.[6,7] At present, research is focused mainly on degeneration of the adjacent intervertebral disk,[8] which is diagnosed radiologically. However, this is a surrogate outcome, and its clinical importance is unknown. The causative relation to surgery is also subject of debate.

Since implants are costly, the discussion should start with the question whether implants are needed in case of surgical therapy for single-level disease. This is the first study that investigates whether the patient-reported clinical outcome differed between patients who were treated by either ACD, ACDF with cage stand-alone, or ACDA.

Methods

Study design and oversight

The guidelines of the CONSORT 2010 statement were followed[9].

Registration of the trial in the registry was done shortly after the starting recruitment of patients, since the authors (at that time) were not aware of the fact that registration in an international register was also necessary since it was already registered in a national register.

The authors confirm that all ongoing and related trials for this drug/intervention are registered. Patients were enrolled between October 5th, 2003 and June 10th 2010 in a single center (the Canisius Wilhelmina Hospital, Department of Neurosurgery, Nijmegen, the Netherlands) in a randomized controlled trial.[10]

Patients were eligible if they suffered from a radicular syndrome in the arm due to one-level cervical degenerative disease of an intervertebral disk at MRI and the involved level was still mobile at dynamic radiographs. They were assigned to surgical treatment consisting of cervical anterior discectomy followed by one of the following three surgical options: fusion by cage stand-alone (ACDF), arthroplasty (ACDA), or no implant at all (ACD). After written informed consent, patients were randomized using a closed-envelope system delivered by an independent co-worker of the medical administrative subdivision of the department.

The trial design was a prospective, double blind, single center randomized study with a three arm parallel group design. The experimental group was ACDA, whereas ACDF and ACD were control group. In the final analysis, it was considered as a superiority design. The type of randomization was 1:1:1. The evening before surgery, the treating surgeon was informed to which group the patient had been allocated.

Although designed as a multiple centre study, the commitments from other centers to contribute were not fulfilled due to several reasons. Reasons were the introduction of more promising and less technically demanding implants on the market, and the lack of financial support.

Although a formal interim analysis was not planned nor made, because it was expected that adding a different implant to a very common procedure would not result in a dramatic positive or negative effect that would justify terminating the trial, whereas the sample size would increase, the inclusion ended before reaching the predefined sample size. After the publication of a meta-analysis indicating that a clinical difference was not present comparing cervical anterior discectomy with fusion and with arthroplasty.[11], we could not justify the continuation of the trial, because the costs for the disk prosthesis were five times higher than a cage stand-alone.

The primary outcome measure was the Neck Disability Index (NDI) at five years. Due to the longer inclusion time and our interest in the log term results, we decided after internal decision to send all patients questionnaires about the NDI in order to have a better impression of the effect of surgery after longer follow-up for the primary outcome measurement. This was done at the moment the last patient completed the five year follow-up. All patients were contacted again and asked to complete the NDI. The last follow-up for all patients was December 1st, 2015.

Secondary outcome measurements were McGill Pain Questionnaire Dutch language version (MPQ-DLV), numeric rating scale (NRS) arm and NRS neck, physical-component summary (PCS) and mental-component summary (MCS) of the 36-item Short-Form Health Survey (SF-36), complications, and reoperations.

Postoperative data collection started 6 weeks postoperatively, and patients were followed for 5 years. At the last follow-up, patients were also asked to complete NRS arm and NRS neck questionnaires estimating the pain during the previous 24 hours.

Patients

All adult patients aged between 18 and 55 years with monoradicular signs and/or symptoms in the arm due to a herniated cervical intervertebral disk and/or an osteophyte at MRI without a history of any cervical spine surgery were eligible. In the original protocol a maximum age was written of 50. We assume that this was a miswriting since in the trial registration the maximum age of was 55. In the subsequent publication[10] the maximum age was also set at 55. The radiological findings should be in accordance with the clinical presentation, and the involved level should be mobile at dynamic radiographs of the cervical spine.

Patients were screened for eligibility after referral. The surgeon offered the possibility of participation to the trial. After at least 48 hours, the patients were contacted again and asked for informed consent.

Patient involvement

We did not involve patients or lay people in the design of the study, since the basic cervical anterior discectomy was already a well-known and accepted treatment. We were interested if adding an implant would be of benefit for the patient. For these purposes we have chosen outcome measures that were known to reflect the clinical situation and daily burden of a patient regarding disability, pain and quality of life. To investigate the benefit we were interested in outcome measurements reported by the patients themselves. Before the study, we extensively studied the burden of participating to the trial for the patients since it was also part of the approval procedure for the ethical committee.
During the study, however, the method for follow-up has significantly been changed since patients requested not to be obliged to visit the outpatient clinic in order to complete the scheduled follow-up moments if they had no complaints. They were willing to complete the questionnaires at home and return them by flat mail.

Interventions

All patients underwent a standard anterior cervical discectomy with bilateral decompression of the nerve roots. If the patients were allocated to ACD, the wound was closed; if they were allocated to ACDF, a cage stand-alone filled with autologous bone was implanted (Brantigan cervical I/F cage, DePuy Spine, Inc., Raynham, MA, USA), and in case of ACDA, a Bryan disk prosthesis (Bryan disk prosthesis, Medtronic, Memphis, TN, USA) was implanted according to the guidelines provided by the company. Postoperatively, none of the patients was prescribed a collar. To prevent heterotopic ossification, only the patients allocated to ACDA were prescribed meloxicam for 2 weeks.

All (four) trial surgeons were senior spinal surgeons experienced in the three types of intervention. Institutional review board approval was obtained (The Ethics Committee CMO Arnhem-Nijmegen, CMO-nr: 2002/200; date of approval May 14, 2003). There was no industry funding.

Outcome measures

The primary outcome was a change in NDI score (scale 0 to 50 points) at the last follow up in November 2015. The NDI is a well known and validated (in multiple languages) outcome-measurement instrument to assess self-rated disability in patients with neck pain.[12-14]

Secondary outcomes were the MPQ-DLV, SF-36, complications, re-operations, and visits to physicians or therapists concerning neck problems after the index surgery for advice or conservative treatment.

MPQ-DLV is a questionnaire that includes several domains. At the moment of completion, the MPQ-DLV and the visual analog scale (VAS) should be rated; whether the complaints were minimal or maximal should also be indicated. A description of the pain should also be given, chosen from a list of adjectives. The number of adjectives was counted (number of word chosen-total [NWC-T]), as well as the sum of the ranks belonging to each adjective (pain rating index total [PRI-T]).

Patients were encouraged to complete the SF-36, MPQ-DLV, and NDI questionnaires themselves, or with the assistance of an independent physician assistant, before they visited or contacted their physician. Patients who wanted to participate, but did not want to visit the outpatient clinic, were offered the possibility to complete the questionnaires at home and return them as hard copy by mail. Baseline NDI, SF-36, and MPQ-DLV were derived carried through 60 months. NDI was also derived at the last follow-up, as was the NRS arm and NRS neck. The response rate was expected to decrease at every follow-up visit during the first 2 years, a well-known phenomenon.[15-17]

For the final follow-up, we emphasized the importance of completing the questionnaires and reminded the patients to do so, if they had not responded.

To optimize participation further, we focused at the last follow-up on the primary outcome (NDI) and the NRS arm and NRS neck. We did not use the MPQ-DLV since it had been shown that responsiveness was higher for NRS compared with VAS,[18] which is only a part of the MPQ-DLV. A high correlation has been found between VAS and NRS supporting interchangeable application.[19]

Recently, a good outcome was also defined as NDI ffi7.[20] Apart from comparison of the NDI value among groups, the proportion of good outcome was also evaluated, although it was not included in the original protocol.

Statistical analysis

We changed the noninferiority assumption into a superiority assumption. We justified this change by the fact that we assumed originally that the three methods resulted in at least similar clinical results, and we expected better results for ACDA.

In the original protocol a 20% difference in excellent outcome was the base for the sample size calculation. An excellent outcome, however, was not exactly defined. Since NDI was a primary outcome measurement, the statement about difference in excellent outcome was interpreted as a difference in NDI.

The total sample size using NDI measuring on a numeric scale (0/50) should be a minimum of 243 patients in order to detect a difference of 10 (or 20% if the percentage scale was used ranging from 0 to 100%). The alternative hypothesis considered a difference of 10 points or 20% in the NDI as clinically relevant. While designing this study, information regarding the minimal clinically important difference was lacking, but recent studies confirmed our assumption.[21,22] We estimated a dropout rate of 10% and therefore estimated that the trial needed to include 270 patients. A dropout rate of 10 % was chosen since it was assumed that a certain amount of people would not participate in the trial after inclusion. Since this number varied comparing several trials we arbitrarily have chosen for 10% in order to have the greatest change to perform an analysis on the previously calculated sample size.

For analysis the intention to treat principle was followed. Analyses of the primary outcome were performed including all patients that had completed the questionnaires at their last follow-up. From the two patients that were lost to follow-up, the baseline data were included, as well as the data at their last follow-up moment. Missing data were not imputed. Only the available data were analyzed. Stratification was not applied.

For NDI as the primary outcome measure, analyses were done using a linear mixed fixed effects model with variance components as covariance type and only a random intercept. In this model treatment group (factor), moment of measurement (factor), and baseline score (covariate) were used to explain the dependent variable (NDI). Age (covariate), enrolment time (covariate), surgeon (factor), and gender (factor) were incorporated in the model, in order to correct for possible confounding.

The proportion of patients with a good outcome within each treatment group was also compared. The same technique with similar variables was used for analyses of MPQ-DLV and SF-36 for two years postoperatively and at 5 years. Analyses for NRS arm and neck were set at the last follow up (December 2015).

The NRS arm and NRS neck at the last follow-up between groups were analyzed separately and investigated using the one-way ANOVA method. Two analyses were done: (1) including all patients irrespectively of the intervening treatment simulating a real life situation, and (2) excluding those patients that underwent additional surgery. HG and RB analyzed the data.

Baseline characteristics were compared between groups using chi-square tests or for categorical data, and one-way ANOVA techniques for continuous data. Numeric data are represented by mean value ± standard deviation (SD). Results of the analyses by the mixed model are represented as mean, standard error and 95% confidence interval (95% CI). For the statistical analyses, SAS version 9.2 (SAS Institute Inc. Cary NC, USA) was used.

Results

Patients

Overall, 272 patients were eligible for inclusion in the trial after screening. However, 18 patients explicitly had a clear preference for one treatment, and 112 refused to participate. Finally, 142 patients gave informed consent and were randomized. The mean age of the study population was 44.9±6.5 years; 50% were female. Baseline characteristics are presented in Table 1, and the distribution of the operated level is presented in Table 2. For 140 (98.6%) patients, the median last follow-up was 8.9±1.9 years (range 5.3 to 12.2 years).

Table 1. *Baseline Characteristics of Included Patients Allocated to Anterior Cervical Discectomy Without Any Implant (ACD), Anterior Cervical Discectomy With Fusion by Cage Stand-Alone (ACDF), or Anterior Cervical Discectomy With Arthroplasty (ACDA) (All characteristics were similar between groups without reaching statistical significance for any difference). Numerical data represented as mean ± SD.*

	ACD	ACDF	ACDA
Age – yr	44.3±5.6	43.1±7.5	44.1±6.4
Gender (F/M)	23/22	22/25	26/24
Smoking (Y/N)	16/29	24/23	27/23
Alcohol consumption (Y/N)	20/25	24/23	27/23
NDI	17.1±6.4	18.8±7.4	18.8±7.5
SF-36 PCS	43.6±12.3	44.0±11.0	44.1±13.9
SF-36 MCS	62.1±18.8	55.7±21.1	58.3±22.2
VAS minimum	21.9±19.2	26.9±21.9	30.1±23.8
VAS maximum	71.6±26.6	68.0±29.1	66.4±29.9
VAS moment	41.9±25.4	39.5±26.0	47.6±29.6
NWC-T	10.5±4.7	8.6±4.8	8.1±4.8
PRI-T	18.5±9.3	15.2±10.2	14.7±10.9
Total	45	47	50

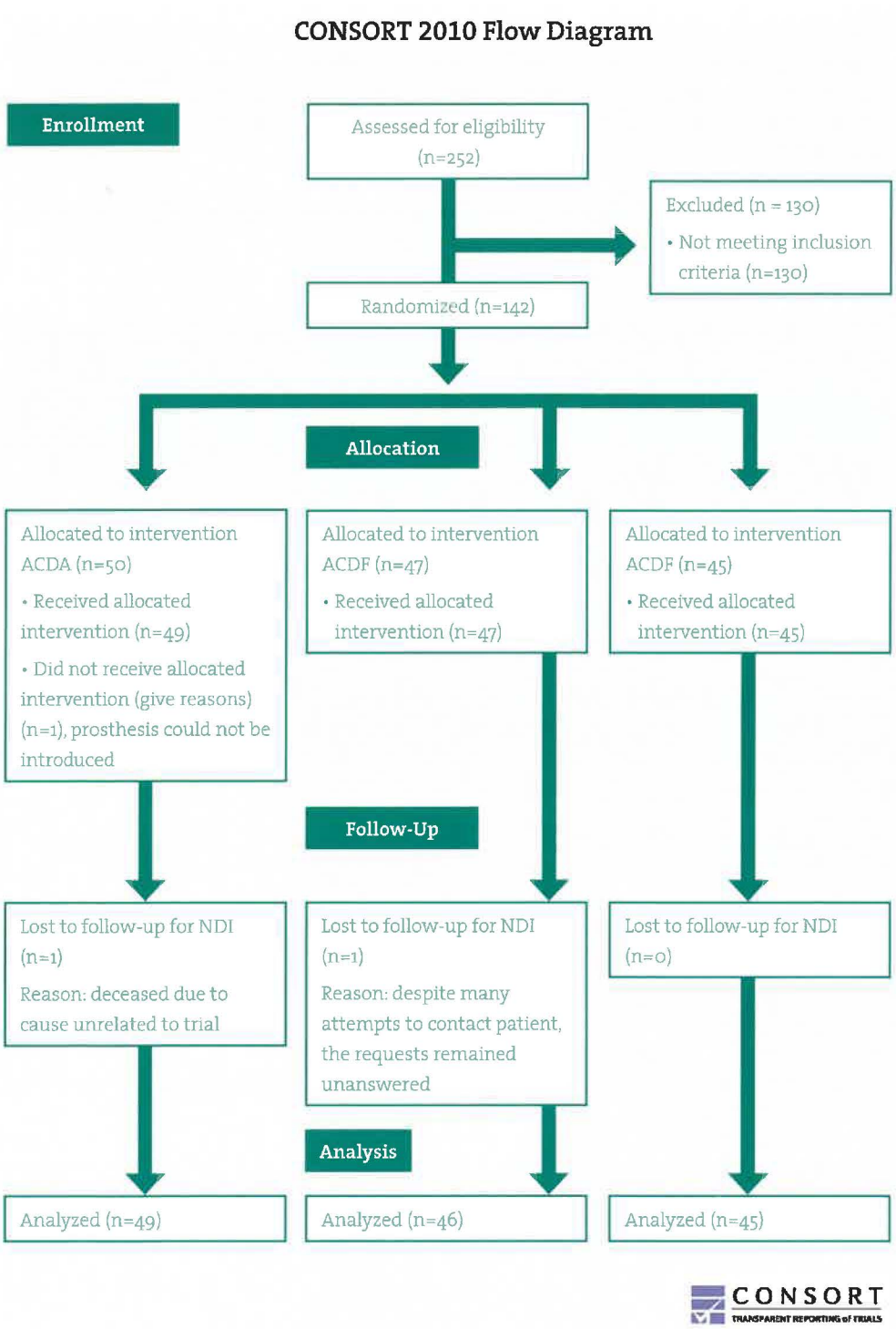
Table 2. *Surgical Level in Relation to Procedure (ACD, ACDF, or ACDA) Statistical difference was not reached (P=0.232).*

Level	Total	ACD	ACDF	ACDA
C4C5	3	1	2	0
C5C6	66	26	19	21
C6C7	73	18	26	29
Total	142	45	47	50

Fifty patients were allocated to ACDA, 47 to ACDF, and 45 to ACD. No differences regarding baseline characteristics were present between the treatment groups. One patient allocated to ACDA died due to a cause unrelated to the intervention, and one patient allocated to ACDF refused to complete the last questionnaire.

One patient was allocated to the arthroplasty group, but intraoperatively it was not possible to introduce the disk prosthesis and a cage was implanted instead. According to the intention-to-treat principle, this patient remained within the arthroplasty group for analysis. The flow diagram according to Consort is represented in Fig 1.

Fig 1: flow diagram according to Consort



Primary outcome

The outcome was dependent upon the baseline score (P = 0.009). Gender, surgeon, time to enrolment and age did not affect the outcome between treatment groups. At two years the primary outcome NDI improved 13.4 ± 0.8 points compared to baseline. This difference was statistically significant (P = 0.009). A statistically significant difference between the three groups was absent (Table 3, Figs 2 and 3)

Table 3. *Estimated Marginal Mean Values of NDI at Different Follow-Up Intervals, based on the linear mixed model computed for baseline NDI score of 18.75**

Postoperative Follow-Up	Mean	Standard Error	95% Confidence Interval		Number of patients		
			Lower Bound	Upper Bound	ACD	ACDF	ACDA
6 weeks	9.2	0.846	7.6	10.9	32	34	36
3 months	7.7	0.846	6.0	9.4	31	36	39
1 year	6.5	0.858	4.9	8.2	30	34	35
2 years	5.5	0.958	3.6	7.4	19	19	24
3 years	7.1	1.046	5.0	9.2	12	13	18
5 years	6.0	1.259	3.5	8.5	4	9	10
9 years	7.5	0.829	5.8	9.1	45	46	49

* At mixed models with fixed effects, only the difference between the preoperative NDI and at 6 weeks' follow-up reached statistical significance. During the remaining follow-up, it remained stable.

Fig 2: NDI with 95%CI at different follow-up moments for the complete sample.

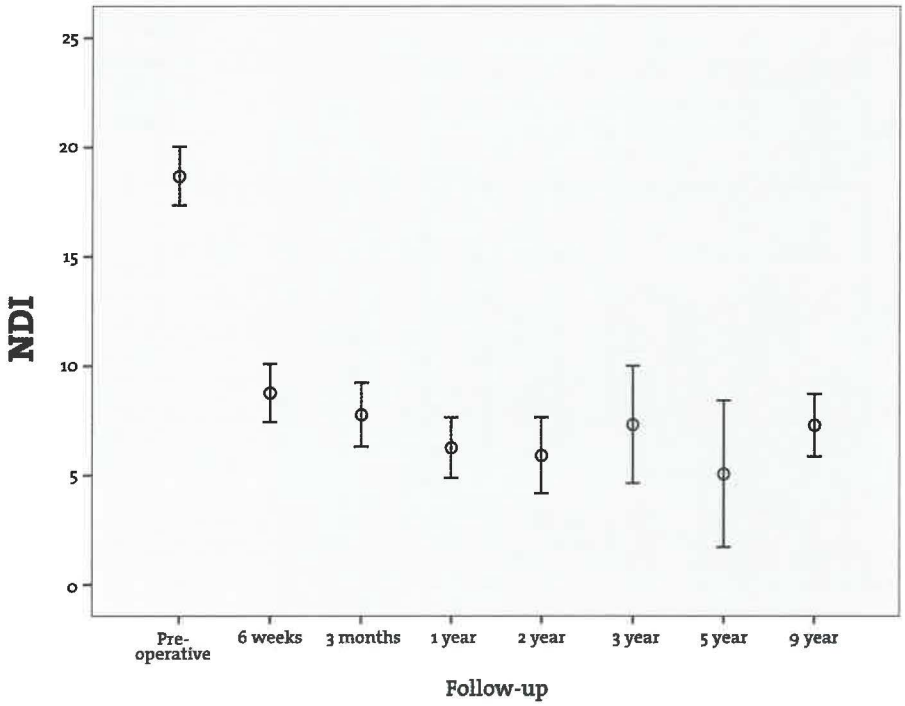
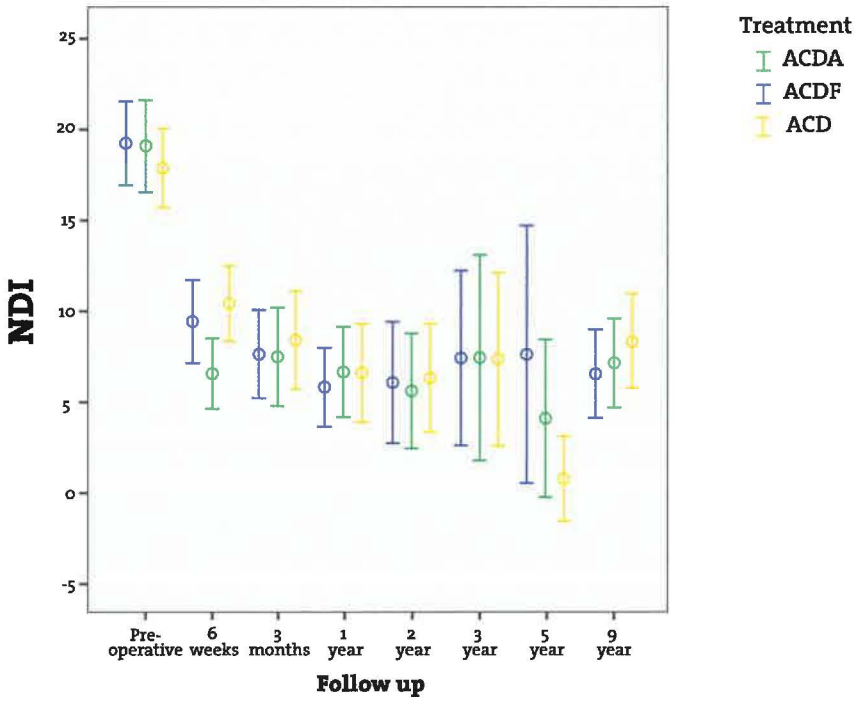


Fig 3: NDI with 95% CI at each follow-up moment and per treatment modality.



At the last follow-up and compared with baseline NDI improved to 7.5 ± 8.5. A statistically significant difference between the groups was absent (P = 0.324). The only clear and statistically significant improvement in NDI was seen between the measurements pre-operatively and 6 weeks postoperatively. Afterward, a clinically relevant change of NDI did occur anymore during follow-up. The mean difference of NDI between 2 years postoperatively and the last follow up was 2.0 ± 0.7 (p = 0.009). Between the treatment groups a statistically significant difference did not exist (Table 4)

Table 4 *The difference in estimated marginal means between groups computed with linear mixed model*

Treatment pair		Mean difference	SE	df	P	Lower bound 95%CI	Upper bound 95%CI
ACDA	ACDF	-0.003	1.308	110.793	0.998	-2.595	2.590
ACDA	ACD	-1.659	1.394	108.981	0.237	-4.422	1.104
ACDF	ACD	-1.656	1.429	109.095	0.249	-4.489	1.176

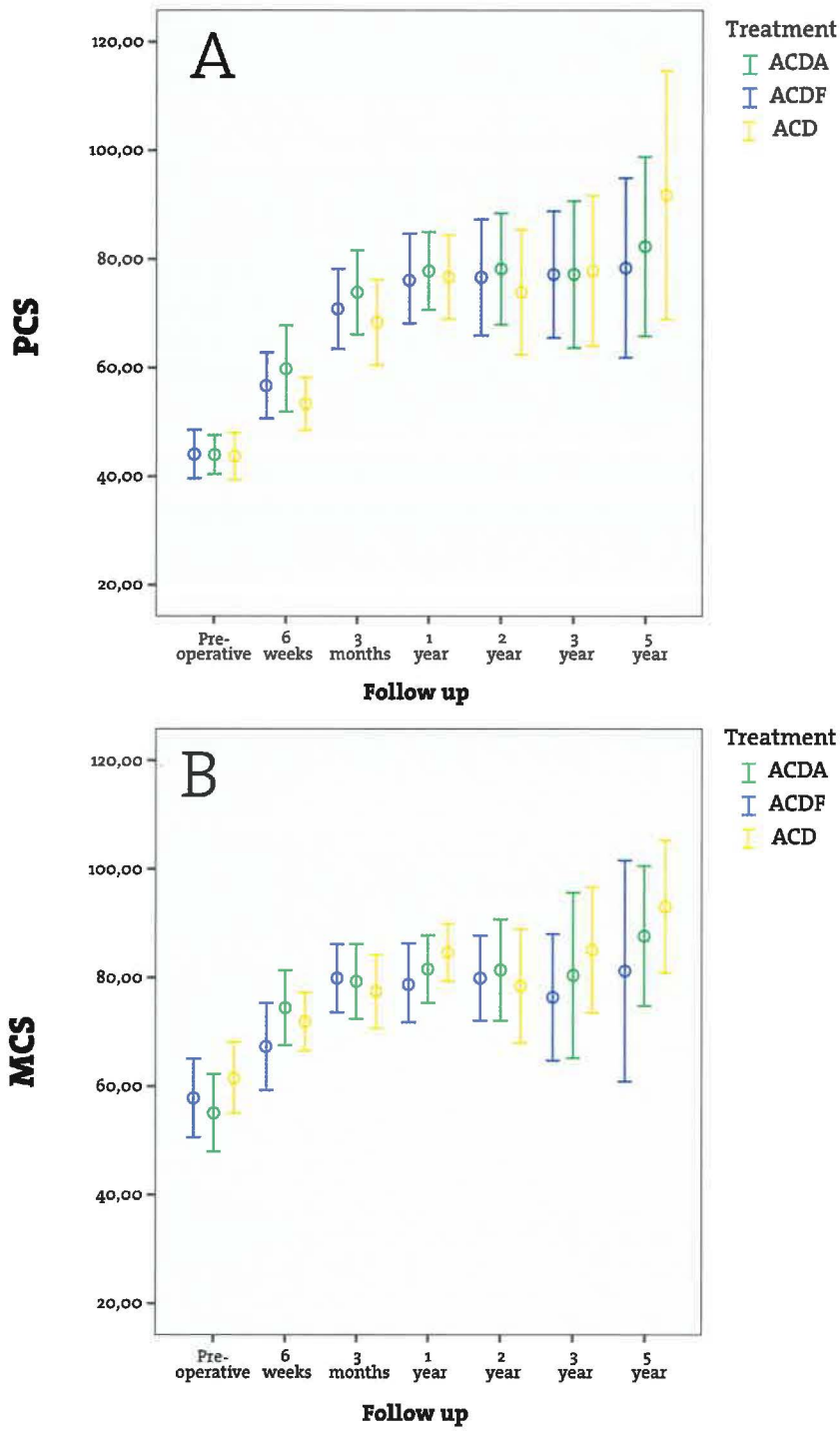
If the patients who underwent a surgical procedure after the index procedure (n = 11) were not incorporated in the analysis, NDI improved to 6.5±7.9, without any statistically significant difference between the treatment modalities (P = 0.832). Of the patients treated by ACDA, 73.5% had a good outcome as defined by a score of NDI ffi7, by ACDF 60.9%, and ACD 57.8% at their last follow-up, but this was not statistically significant (P = 0.239).

The enrolment time to model time[23] did not affect the outcome and was, therefore not included in the model. Gender, age, and surgeon were not related to any clinical relevant outcome measurement.

Secondary outcomes

Regarding the summary scales of SF-36 at two years the mean improvement of PCS was 32.1±2.5. A statistical difference between the treatment modalities was not found at any follow-up moment (P = 0.873). MCS improved on average 22.8±2.1 without any statistically significant difference between the groups at any follow-up moment (P = 0.874) (Fig 4).

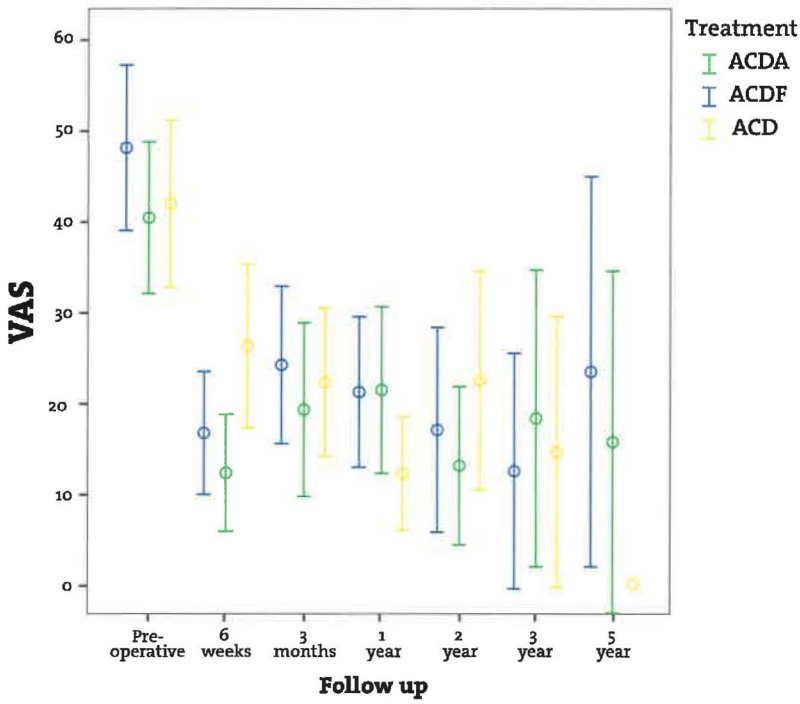
Fig 4: *graph depicting PCS (A) and MCS (B) with 95% CIs at different follow up and for each treatment.*



VAS as part of the MPQ-DLV improved to 17.3±24.0, 9.2±16.5 when the complaints were minimal, and 24.4±31.4 if maximal. Only the VAS at the moment of completing the questionnaire is shown in Fig 5. All others had a similar pattern to NWC-T and PRI-T. NWC-T was 4±5, and PRI-T was 6.3±9.7. None of them reached statistical significance between the treatment modalities (VAS at moment completing questionnaire: P = 0.429, VAS minimal pain: P = 0.534; VAS maximal P = 0.593; NWC-T: P = 0.690; PRI-T: P = 0.657).

MPQ-DLV was completed by only 22 patients at 5 years. Therefore, we represent them only in the graphs (Fig 5). Instead, we mention the NRS arm and NRS neck, which were obtained at the last follow-up in 140 patients. NRS arm was 1.8±2.5, and NRS neck was 1.9±2.6. A statistical difference between the treatment groups was not present (P = 0.622 and 0.496, respectively).

Fig 5: *VAS at the moment of completing Questionnaire with 95% CI at different follow-Up moments until 5 years*



Complications

In 13 patients (9.2%), complications occurred that were not related to signs or symptoms of recurrent compression or nerve root involvement at the adjacent segment (Table 4). Urinary tract infections, pulmonary infections, deep venous thrombosis, pulmonary embolism, or deep wound infections did not occur. A superficial wound infection was present in one (0.7%) patient, hoarseness was reported in four (2.8%) patients, dysphagia in seven (4.9%) patients, and a post-operative haemorrhage warranting surgical re-exploration in one (0.7%) patient (Table 5).

Table 5. *Complications Related to Treatment Group (Number/Percentage of Group).**

Complication	ACD	ACDF	ACDA
Number of patients	45	47	50
Urinary tract infection, n (%)	0 (0)	0 (0)	0 (0)
Pulmonary infection, n (%)	0 (0)	0 (0)	0 (0)
Deep venous thrombosis/ pulmonary embolism, n (%)	0 (0)	0 (0)	0 (0)
Superficial wound infection, n (%)	0 (0)	1 (2.1)	0 (0)
Deep wound infection, n (%)	0 (0)	0 (0)	0 (0)
Hoarseness, n (%)	3 (6.7)	1 (2.1)	0 (0)
Dysphagia, n (%)	1 (2.2)	4 (8.5)	2 (4.0)
Postoperative hemorrhage, n (%)	1 (2.2)	0 (0)	0 (0)
Total, n (%)	5 (11.1)	6 (12.8)	2 (4.0)

Recurrent nerve root symptomatology

Eleven (7.8%) patients underwent surgery due to recurrent signs or symptoms related to compression of a nerve root at the index level in three (2.1%) patients and at the adjacent level in eight (5.6%) patients (Table 6). The difference between groups did not reach statistical significance (P = 0.132)

Table 6. *Surgery for Recurrent Signs and Symptoms due to Nerve Root Compression at the Index Level or Adjacent Segment.*

Procedure	ACD	ACDF	ACDA
Surgery for adjacent segment disease, n (%)	3 (6.7)	5 (10.6)	0 (0)
Surgery for recurrent compression at index level, n (%)*	1† (2.2)	1 (2.1)	1 (2.0)
Posterior surgery	1 (2.2)	1 (2.1)	1‡ (2.0)
Anterior surgery	1 (2.2)	0(0)	0 (0)
Total, n	4	6	1

* Approaches for the surgery at the index level is subdivided in anterior or posterior approach.
† One patient was also operated anteriorly because of insufficient result of the first posterior re-exploration.
‡ One patient visited the outpatient clinic for recurrent signs and symptoms before completing the NDI questionnaire. This crossed the radiological examinations, after which she was offered surgical therapy for recurrent stenosis at the index level. She was not included in this analysis.

Consultation of physicians or therapists

Of 140 patients at the last follow-up, 58 (41.4%) consulted at least once a physician or therapist after the index surgery for problems relating to their neck. The number of patients was equally divided among treatment allocations (ACDA 19, ACDF 19, ACD 20) without any statistically significant difference (P = 0.872). The consulted caregivers (with the number of patients who visited the caregiver in parentheses): physiotherapist (26); chiropractor (2); osteopath (1); neurosurgeon (15); orthopedic surgeon (5); general physician (1); pain consultant (6); and neurologist (5). Some patients consulted more than one caregiver. A difference between treatment groups did not exist (P = 0.144).

Discussion

This study is unique since three surgical options were compared: ACD, ACDF with cage stand-alone, and ACDA. The follow-up period is the longest in literature, and the response rate for primary outcome NDI is very high (98.6%).

For the first time, the clinical outcome of ACDA is compared with ACDF with cage stand-alone or ACD. Until now, arthroplasty has been compared only with ACD with fusion with plate fixation. Comparing ACD with arthroplasty and ACD with fusion by plate is comparing two different surgical methods, since the dissection is wider and slightly different in case of the implant of a cage. ACDF with cage stand-alone will resemble more the technique of ACDA.[24] In the current trial, the only difference is whether an implant is chosen and, if so, which implant.

It is remarkable that ACDF is considered as the gold standard, since sound evidence is still lacking in literature. In one study comparing four groups, a statistically significant difference regarding in favor of an additional implant was found after short-term follow-up.[25] The authors of that study commented on the findings that the results were flawed by the small sample size of 125 patients in total.[25] A recent systematic review did not show any clinical superiority of ACDF.[26]

Patients want a treatment that provides them with long-term relief of signs and symptoms, so focus on clinical outcome is important. Therefore, we did not remove from our analysis the patients who have been operated on after the index surgery. This analysis shows clearly what the result of a treatment is after 9 years, including intervening surgical therapies. We did not focus on radiological outcome. In our opinion, surrogate outcomes such as radiological deterioration of adjacent levels without any clinical sign or symptom are not relevant for patients.

In the end, comparing ACD with ACDF and ACDA, the clinical result is similar. Irrespective of the treatment, there is a small change indicating that an additional surgical procedure is needed. Although not statistically significant, it seems that surgery for adjacent segment disease is less often provided for ACDA. Although proponents of the use of cervical disk prostheses claim that prevention of adjacent segment disease is their major benefit compared with ACDF with plate fixation, meta-analyses still show contradicting results.[27-29]

Although our results will contribute to future meta-analyses on this topic, we do not feel confident to recommend disk prostheses as a standard option. Health economics should be considered. Since clinical outcome is not involved in the end, societal costs and hospital costs are involved. These differ between countries. Furthermore, it is important to calculate the number to treat to prevent one extra patient from developing adjacent segment disease. If the costs of an implant are relatively very high (as are disk prostheses in the Netherlands), it might not be economically worthwhile to advise disk prosthesis. A thorough economic evaluation is warranted.

We will not advocate new studies including new patients. One possibility would be to collect all individual patient data in the numerous randomized controlled trials that have been performed comparing ACDA with ACDF by fixation with plate by an independent researcher who has no relation to the industry and does not favor one method above another. Focus should be on the difference in proportion of good outcome. Whether disk prosthesis should be advised so as to prevent adjacent segment disease cannot be concluded based on the results of this study. Cost analysis in relation to number to treat is important.

Limitations

Ending the trial before reaching the calculated sample size might be explained as a major flaw. As explained in the Methods section, we could not justify continuation of inclusion. We felt a major obligation to follow the patients and report on the results. Given the presented results following the protocol, in our opinion, the conclusion would be the same as when the calculated sample size was achieved. Because of these sample size constraints, the risk of not detecting even modest changes is still present, but nevertheless we would like to describe the findings

as inconclusive.[30] However, recent insights about the definition of a good outcome[20] make us doubtful of the correctness of this decision. Comparing the proportions of good outcome, extending the trial could probably have led to a more conclusive statement about a difference in treatments.

The long time to include patients might also be addressed as a flaw. However, since the start of the trial, operative techniques have not been changed, and this will not have influenced the results. Although the trial was developed initially as a multi-center trial, other initially supporting centers did not participate. The mono-center execution of the trial might be defined as a flaw as well, but in our opinion, this will not affect generalizability. The inclusion criteria were very clear, the operative method is not exclusive, the decision to offer new surgery was also clearly described, and the primary outcome NDI is a patient-reported outcome that has been validated worldwide.

We think that this small deviation of the original protocol by adapting the maximum age did not influence the results since movement of the involved disc level at the dynamic X Rays was required for inclusion.

Not reporting the short-term outcomes at 3 months and 1 year did not influence the interpretation of the results, since after 6 weeks postoperatively the results did not alter. This is in accordance with our experience in daily clinical practice. Major improvements in the clinical situation are not expected anymore after the first postoperative out patient clinical visit (approximately six weeks postoperatively).

Another shortcoming is the decision to make the patient responsible for completing the questionnaires without strict control. Therefore, the response gradually decreased during follow-up. However, at the last follow-up, the response rate was nearly 100%, probably because of a more active attitude from the researchers. This contributed to a good response regarding primary outcome NDI, NRS, and reoperations. We are aware that a VAS measured in millimeters is not similar to the NRS in scale 0 to 10, but comparison of severity of pain is still possible [19]. Therefore, the VAS as part of the secondary outcome MPQ-DLV can be compared with the NRS arm and neck. These outcomes are relevant to patients. The effect of blinding is always a subject of discussion. However, since patients always see their postoperative radiograph, blinding was not possible. Since we advocated the trial to the patients by emphasizing that we really did not know what the best treatment was, the effect of not blinding will be minimal. Further, one might criticize the decision to analyze the whole group including those patients with additional surgery after the index surgery. In our opinion, though, this resembles the daily clinical practice, since patients are interested in the final result after a certain treatment including additional operative or non-operative treatments. This might result in worse treatment results because of complaints occurring because of a second operated level. Therefore, it could be considered as a pragmatic solution, since we do not optimize our analysis only in order to determine efficacy.[17]

Selection bias based on failure for concealment might have occurred. However, since the surgeons did not have any preference for a method, the patients were included after they had the

possibility to consider participation, and the allocation was known the evening before surgery we think the change for this kind of bias is minimized. In fact, none of the patients refused the allocated type surgery, nor did the surgeons refuse to perform it.

Conclusion

This randomized trial could not detect a difference between three surgical modalities for treating a single-level degenerative disk disease. Anterior cervical discectomy without implant seems to be similar to anterior cervical discectomy with fusion by cage stand-alone or with disk prosthesis. Due to the small study sample size, this statement should be considered as inconclusive so far. Although a difference was noted in the incidence of adjacent disc disease a definitive conclusion can not be made due to the small sample size.

Acknowledgments

We express our gratitude to all patients that participated in this trial. It is obvious that without their support we could not have made this contribution to medical knowledge. We want also thank Gary Anderton from Radboud in'to Languages for revising style and English grammar.

References

1. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT. Epidemiology of cervical radiculopathy. A population-based study from Rochester, Minnesota, 1976 through 1990. *Brain.* Apr 1994;117 (Pt 2):325-335.

2. Schoenfeld AJ, George AA, Bader JO, Caram PM, Jr. Incidence and epidemiology of cervical radiculopathy in the United States military: 2000 to 2009. *J Spinal Disord Tech.* Feb 2012;25(1):17-22.

3. Kuijper B, Tans JT, Beelen A, Nollet F, de Visser M. Cervical collar or physiotherapy versus wait and see policy for recent onset cervical radiculopathy: randomised trial. *BMJ.* 2009;339:b3883.

4. Woods BI, Hilibrand AS. Cervical radiculopathy: epidemiology, etiology, diagnosis, and treatment. *J Spinal Disord Tech.* Jun 2015;28(5):E251-259.

5. Marawar S, Girardi FP, Sama AA, et al. National trends in anterior cervical fusion procedures. *Spine (Phila Pa 1976).* Jul 1 2010;35(15):1454-1459.

6. Hu Y, Lv G, Ren S, Johansen D. Mid- to Long-Term Outcomes of Cervical Disc Arthroplasty versus Anterior Cervical Discectomy and Fusion for Treatment of Symptomatic Cervical Disc Disease: A Systematic Review and Meta-Analysis of Eight Prospective Randomized Controlled Trials. *PLoS One.* 2016;11(2):e0149312.

7. Zhu Y, Tian Z, Zhu B, Zhang W, Li Y, Zhu Q. Bryan Cervical Disc Arthroplasty Versus Anterior Cervical Discectomy and Fusion for Treatment of Cervical Disc Diseases: A Meta-analysis of Prospective, Randomized Controlled Trials. *Spine (Phila Pa 1976).* Jun 2016;41(12):E733-741.

8. Zhu Y, Zhang B, Liu H, Wu Y, Zhu Q. Cervical Disc Arthroplasty Versus Anterior Cervical Discectomy and Fusion for Incidence of Symptomatic Adjacent Segment Disease: A Meta-Analysis of Prospective Randomized Controlled Trials. *Spine (Phila Pa 1976).* Oct 1 2016;41(19):1493-1502.

9. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med.* Mar 24 2010;7(3):e1000251.

10. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord.* 2006;7:85.

11. Bartels RH, Donk R, Verbeek AL. No justification for cervical disk prostheses in clinical practice: a meta-analysis of randomized controlled trials. *Neurosurgery.* Jun 2010;66(6):1153-1160; discussion 1160.

12. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring the quality and effectiveness of cervical spine surgery in registry efforts: determining the most valid and responsive instruments. *Spine J.* Jun 1 2015;15(6):1203-1209.

13. MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther.* May 2009;39(5):400-417.

14. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther.* Sep 2008;31(7):491-502.

15. Latza U, Stang A, Bergmann M, et al. [The problem of response in epidemiological studies in Germany (part I)]. *Gesundheitswesen.* May 2004;66(5):326-336.

16. Hazen RJ. Population based control selection and nonresponse among case-control studies in cancer research. *Ann Epidem.* 2005;15(8):640.

17. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med.* Aug 4 2016;375(5):454-463.

18. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* Oct 2011;152(10):2399-2404.

19. Bahreini M, Jalili M, Moradi-Lakeh M. A comparison of three self-report pain scales in adults with acute pain. *J Emerg Med.* Jan 2015;48(1):10-18.

20. Donk R, Verbeek A, Verhagen W, Groenewoud H, Hosman A, Bartels R. The Qualification of Outcome after Cervical Spine Surgery by Patients Compared to the Neck Disability Index. *PLoS One.* 2016;11(8):e0161593.

21. Auffinger BM, Lall RR, Dahdaleh NS, et al. Measuring surgical outcomes in cervical spondylotic myelopathy patients undergoing anterior cervical discectomy and fusion: assessment of minimum clinically important difference. *PLoS One.* 2013;8(6):e67408.

22. Lauche R, Langhorst J, Dobos GJ, Cramer H. Clinically meaningful differences in pain, disability and quality of life for chronic nonspecific neck pain - a reanalysis of 4 randomized controlled trials of cupping therapy. *Complement Ther Med.* Aug 2013;21(4):342-347.

23. Altman DG, Royston JP. The hidden effect of time. *Stat Med.* Jun 1988;7(6):629-637.

24. Ahn SS, Paik HK, Chin DK, Kim SH, Kim DW, Ku MG. The Fate of Adjacent Segments After Anterior Cervical Discectomy and Fusion: The Influence of an Anterior Plate System. *World Neurosurg.* May 2016;89:42-50.

25. Barlocher CB, Barth A, Krauss JK, Binggeli R, Seiler RW. Comparative evaluation of microdiscectomy only, autograft fusion, polymethylmethacrylate interposition, and threaded titanium cage fusion for treatment of single-level cervical disc disease: a prospective randomized study in 125 patients. *Neurosurg Focus.* Jan 15 2002;12(1):E4.

26. Botelho RV, Dos Santos Buscariolli Y, de Barros Vasconcelos Fernandes Serra MV, Bellini MN, Bernardo WM. The choice of the best surgery after single level anterior cervical spine discectomy: a systematic review. *Open Orthop J.* 2012;6:121-128.

27. Luo J, Gong M, Huang S, Yu T, Zou X. Incidence of adjacent segment degeneration in cervical disc arthroplasty versus anterior cervical decompression and fusion meta-analysis of prospective studies. *Arch Orthop Trauma Surg.* Feb 2015;135(2):155-160.

28. Verma K, Gandhi SD, Maltenfort M, et al. Rate of adjacent segment disease in cervical disc arthroplasty versus single-level fusion: meta-analysis of prospective studies. *Spine (Phila Pa 1976).* Dec 15 2013;38(26):2253-2257.

29. Botelho RV, Moraes OJ, Fernandes GA, Buscariolli Ydos S, Bernardo WM. A systematic review of randomized trials on the effect of cervical disc arthroplasty on reducing adjacent-level degeneration. *Neurosurg Focus.* Jun 2010;28(6):E5.

30. Pocock SJ, Stone GW. The Primary Outcome Fails - What Next? *N Engl J Med.* Sep 1 2016;375(9):861-870.

Chapter 5

Factors determining a good recovery after surgery for symptomatic single level cervical degenerative disc disease: a prospective cohort study.

Roland Donk, André Verbeek, Wim Verhagen, Hans Groenewoud,
Allard Hosman, Ronald Bartels

Abstract

Background

Anterior cervical discectomy is a well-known and frequently performed procedure. Although several variations of the procedure exist, the basics remain similar: decompression of the impinged nerve root and/or compressed spinal cord. To counsel patients properly before an operation, information about the expected outcome is necessary.

Objective

To define factors that predict outcome after an anterior cervical discectomy procedure because of single-level degenerative disease.

Methods

From a prospective cohort of patients, outcome was determined by Neck Disability Index (NDI). An NDI score ffi 7 was considered as a good outcome. Predictive factors were determined through univariate and regression techniques.

Results

140 patients were included for this analysis. Mean follow-up was 9.1 ± 1.9 years (5.6–16.0 years). Predictive factors for a less than good outcome were smoking behavior (odds ratio [OR] 4.318, 95% confidence interval [CI]: 1.703–10.948), index level at c5c6 (OR 4.250, 95% CI: 1.675–10.781), and not having been operated for adjacent segment disease (OR 9.615, 95% CI: 1.003–90.909). This was also true for a positive increase of the postoperative sagittal angle at the index level (OR 1.095, 95%CI: 1,013–1,183). The preoperative NDI score strongly correlated with the NDI score at last follow-up.

Conclusion

Factors that strongly predict a suboptimal outcome are smoking and , very remarkable, surgery at c5c6. Furthermore, a positive increase of the sagittal angle at the index level also contributed to a good outcome, as well as not having an operation due to adjacent disc disease.

Introduction

Cervical degenerative disease with radicular symptoms is encountered frequently in daily practice. In general, a conservative treatment is instituted with good results. However, if the treatment fails, surgery is a valid option. The outcome of surgery depends in part upon the expectations of the patient. Therefore, preoperative consultation mentioning pros and cons of surgery is extremely important.

To provide optimal information knowledge of predictive factors is helpful. Several studies tried to identify factors that were predictive of a good outcome.[1-5] However, they used scales such as the Neck Disability Index (NDI) or Visual Analog Scale (VAS) that are continuous. Predictive factors were related to these respective scores by regression techniques, by relating those factors to the minimal clinically important difference (MCID) and other indirect methods to address the amount of patient satisfaction. Recently, a clear cut-off value of the NDI has been described for a good outcome according to the rating of the patient.[6] Based on this value of the NDI, a patient-reported outcome measurement tool, a clear distinction can be made between a good and less than good outcome. In this study, we evaluate factors that could be a possible predictive factor of a good outcome after surgery.

Methods

The Ethics Committee CMO Arnhem–Nijmegen approved the trial (CMO-nr, 2002/200). The study was carried out in accordance with the World Medical Association Declaration of Helsinki.[7] The STROBE statement was followed.[8]

Patients participated after informed consent in a randomized controlled trial investigating whether a difference existed in outcome between anterior cervical discectomy without fusion (ACD), with fusion by stand-alone cage (ACDF), or with implantation of a disc prosthesis (ACDA). The results of this trial only focusing on outcome have been reported.[9] Only patients who completed the NDI at the last follow-up were included in the current study. Because of the research question, it was considered a prospective cohort of patients. An NDI score ffi 7 is considered as a good outcome.[6]

The following factors were included in the analysis: age, body mass index (BMI), surgical procedure, level of surgery, sagittal alignment of cervical spine preoperatively (lordosis, straight, kyphosis) assessed by the modified Toyama procedure,[10] surgeon, history of low back surgery, surgery for cervical adjacent segment disease (ASD) after index surgery, sagittal angle of the index level measured by the posterior tangent method 1 day postoperatively,[11] and preoperative NDI score.

For statistical analysis, SPSS version 22.0 (Armonk, NY: IBM Corp.) was used. Categorical data were analyzed by chi-square tests, and continuous data by student-t tests. Continuous variables are represented as mean ± standard deviation (minimum – maximum).

First, a univariate analysis was performed exploring the relation between outcome and the factor. If p ffi 0.10, the variable was included in the next analysis, a logistic regression with backward selection for the selected variables. In this model odds ratios (ORs) were estimated, as well as the 95% confidence intervals (CIs). A p value ffi 0.05 was considered statistically significant.

Results

Of the 142 patients who were included in the trial, 140 could be included for this analysis. Half of them were male. The mean age was 54.2 years ± 6.5 years (18.3–54.2 years) and mean follow-up was 9.1 ± 1.9 years (5.6–16.0 years). Ninety (64.3%) patients reported an NDI score ffi 7 and were considered as having a good result. Univariate testing for the following factors did not resolve statistical significance: surgeon (p = 0.553), surgical procedure (p = 0.213), BMI (p = 0.284), age (p = 0.736), history of previous low back surgery (p = 0.101), and preoperative sagittal alignment (p = 0.447).

The difference in good outcome between males (72.9%) and females (55.7%) reached statistical significance (p = 0.034). Patients who smoked had a worse outcome than those who did not (p = 0.007). The history of smoking was known for 127 (90.7%) patients, and these patients were included in the logistic regression. 43.3% of the female patients smoked, as did 28.1% of the males. This difference did not reach statistical significance (p = 0.085). Six patients were operated within the follow-up operated for a symptomatic adjacent segment disease and perceived a less than good outcome. This contrasted with the group who did not suffer from ASD and had a less than good outcome (33.3%); this difference reached statistical significance (p = 0.017). Outcome seemed to be related to thec4c5, c5c6 and c6c7 levels of surgery (p = 0.032), with a good outcome related to c4c5 and c6c7 (75.0% and 74.3%, respectively), whereas those patients who were operated on at c5c6 level experienced more often a less than good result (good outcome: 53.0%). It should be mentioned that only 4 patients were treated at c4c5 level. The postoperative angle of the index level was 1.9 ± 5.6 degrees in the group that had a good outcome and −0.3 ± 6.7 in the group with a less than good outcome (p = 0.037). This differed statistically significantly among the applied treatment: ACDA 1.5 ± 5.3 degrees, ACDF 4.3 ± 5.9 degrees, and ACD −2.2 ± 5.3 degrees (p < 0.001). The angles within each treatment related to outcome are represented in Table 1. The preoperative NDI score was 17.4 ± 7.3 and 20.0 ± 6.7 for patients with good result and those with less than good result, respectively (p = 0.037).

The following factors were included in the model: gender, smoking, level of index surgery, ASD, postoperative angle of the index level, and initial NDI score. After backward selection, the only factor that was removed was gender (p =0.221). The results are represented in Table 2. A statistically significant difference was seen with smoking (p = 0.002), ASD (p = 0.048), level of surgery comparing c5c6 with c6c7 (p = 0.002), and postoperative angle of index level (p = 0.021).

Discussion

Knowledge of factors that predict long-term outcome is essential in counseling patients. Analysis of predictive factors has been done before.[1-5] However, a good outcome was not defined using

a cut-off value based on the patients' rating. For example, in the study by Hermansen et al., a good outcome was chosen with values corresponding to the minimal clinically important difference (MCID).[5] MCID was developed to characterize the smallest amount of change between the pre- and postoperative situation that the patient perceived as meaningful.[12] It does not automatically correspond to a good outcome as rated by the patient. Peolsson and co-workers did not define a good outcome, but used a statistical approach with continuous scales for NDI, VAS arm, and VAS neck.[4,13]

This is the first report that used a clear definition of good outcome or less than good outcome.[6] Several factors were related with a good outcome after logistic regression with stepwise selection of variables.

The relation between a less than good outcome and smoking has been described earlier.[13,14] The negative effect of smoking on pain in various circumstances is well known.[15-18] Smoking might cause up-regulating of the nicotinic acetylcholine receptors throughout the nervous system, and thereby, non-selectively antagonizing analgesic effects.[19]

Surgery for adjacent-level disease was related to a less than good outcome. Repeated spine surgery for other conditions has also been related to worse outcomes.[20,21] However, caution is warranted while interpreting this factor since the number of cases in this series with ASD is small (only 6 patients).

The interpretation of the predictive value of the angle of the index level is not straightforward. If the sagittal angle increased 1 degree, the probability to have a less than good outcome diminished. It does not automatically mean that each kyphotic spine did not have a good outcome. If the angles were determined within each treatment group, most of the patients with a suboptimal outcome have a positive postoperative angle at the index level. It seemed that especially patients within the ACD group contributed to the negative mean value of the total sample of the patients. This was particularly due to the magnitude of the value of the postoperative sagittal angle in a few patients. Although the mean angle in the group of patients with less than good outcome was negative (−0.3), it can be questioned whether the angle should be defined as clinically kyphotic.

The relation between outcome and the level of surgery is very remarkable. Patients who were operated on because of pathology at c5c6 had an OR > 4 for having a suboptimal outcome. We cannot define any biologically plausible explanation for this finding. Level of surgery was introduced in the model as a known baseline characteristic, and surprisingly it was a factor of importance. We did not find any similar clinical findings in literature perhaps because previously this was not considered relevant and, therefore, not investigated. However, in asymptomatic people, degenerative changes were found most frequently at the c5–6 level, followed by c6–7.[22] In an autopsy study, Friedenberg et al.[21] found that the c5–6 level was most frequently involved in spine degeneration, followed by c6–7 with the c2–3 level least involved. The Luschka's joints were altered markedly in the lower 3 levels, most frequently at the c5–6 region.[23] So maybe there

is a mechanical reason not only for more common degeneration at the c5–6 level, but also for a less favorable outcome.

This study has some biases. Although it was a prospective study, smoking behavior was not noted in all patients. Therefore, for the final regression, 83% of the patients could be used. Nevertheless, we feel comfortable that the results will be consistent in larger series. It is very likely that CIs will become smaller. Since it is known that re-operation might contribute to less good outcomes, we think that in larger series this will become more evident, with CIs moving away from 1.0.

A major strength of this study is the introduction of a cut-off value for NDI that has been proven to relate to a good outcome as rated by the patients.[6] This contributed to a clear and easy to understand interpretation of the results. For counseling a patient and their relatives, it is extremely important to manage the expectations properly to achieve an optimal result. Although these patient fulfilled the inclusion criteria for a randomized controlled trial, we feel confident that these data can be used in clinical practice. However, confirmation of these findings in future studies is obligatory.

Conclusion

In conclusion, an anterior approach for symptomatic cervical single-level disc degeneration might be successful. However, obviously negative predictive factors were smoking and c5c6 as the symptomatic level. Less prominent negative predictive factors were re-operation for adjacent segment disease and, even to a lesser degree, some loss of lordosis. No difference could be established between used anterior techniques or gender.

Table 1. *Postoperative angles (degrees) within each treatment divided among good or less than good outcome of total sample of patients (N = 140). Number of patients per group are represented (n), as well as the mean difference in angle per treatment group and the result of the t-test comparing the angles within a treatment group*

NDI = Neck Disability Index.

	ACDA		ACDF		ACD	
Outcome	n	Angle	n	Angle	n	Angle
NDI ffi 7	36	1.8 ± 5.2	28	5.2 ± 4.9	26	−1.4 ± 4.7
NDI > 7	13	0.4 ± 5.6	17	2.8 ± 7.1	20	−3.3 ± 5.9
Mean difference		1.4 ± 1.7		2.4 ± 1.8		1.9 ± 1.6
p value		0.409		0.187		0.220

Table 2. *Factors included in logistic model. OR related to less than good outcome*

Factor	OR	95% CI
Smoking	4.318	1.703–10.948
Level of index surgery c5c6 compared with c6c7	4.250	1.675–10.781
Postoperative angle	0.913	0.845–0.987
Initial NDI	1.065	1.002–1.133
Not being operated for adjacent segment disease	0.104	0.011–0.977

CI = confidence interval; NDI = Neck Disability Index; OR = odds ratio.

References

1. Anderson PA, Subach BR, Riew KD. Predictors of outcome after anterior cervical discectomy and fusion: a multivariate analysis. Spine (Phila Pa 1976). Jan 15 2009;34(2):161-166.

2. Mayo BC, Massel DH, Bohl DD, et al. Preoperative mental health status may not be predictive of improvements in patient-reported outcomes following an anterior cervical discectomy and fusion. J Neurosurg Spine. Feb 2017;26(2):177-182.

3. Alvin MD, Miller JA, Lubelski D, et al. The Impact of Preoperative Depression and Health State on Quality-of-Life Outcomes after Anterior Cervical Diskectomy and Fusion. Global Spine J. Jun 2016;6(4):306-313.

4. Peolsson A, Vavruch L, Oberg B. Predictive factors for arm pain, neck pain, neck specific disability and health after anterior cervical decompression and fusion. Acta Neurochir (Wien). Feb 2006;148(2):167-173; discussion 173.

5. Hermansen A, Hedlund R, Vavruch L, Peolsson A. Positive predictive factors and subgroup analysis of clinically relevant improvement after anterior cervical decompression and fusion for cervical disc disease: a 10- to 13-year follow-up of a prospective randomized study: clinical article. J Neurosurg Spine. Oct 2013;19(4):403-411.

6. Donk R, Verbeek A, Verhagen W, Groenewoud H, Hosman A, Bartels R. The Qualification of Outcome after Cervical Spine Surgery by Patients Compared to the Neck Disability Index. PLoS One. 2016;11(8):e0161593.

7. Fuson RL, Sherman M, Van Vleet J, Wendt T. The conduct of orthopaedic clinical trials. J Bone Joint Surg Am. Jul 1997;79(7):1089-1098.

8. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. Oct 16 2007;147(8):573-577.

9. Donk R, Verbeek A, Verhagen WIM, Groenewoud H, Hosman A, Bartels R. What's the Best Surgical Treatment for Patients With Cervical Radiculopathy due to Single-Level Degenerative Disease? A Randomized controlled trial. PLoS One. 2017;accepted.

10. Donk RD, Fehlings MG, Verhagen WIM, et al. An assessment of the most reliable method to estimate the sagittal alignment of the cervical spine: analysis of a prospective cohort of 138 cases. J Neurosurg Spine. May 2017;26(5):572-576.

11. Donk R, Arnts H, Verhagen WIM, Groenewoud H, Verbeek A, Bartels R. Cervical sagittal alignment after different anterior discectomy procedures for single level cervical degenerative disc disease: randomized controlled trial. Acta Neurochir (Wien). 2017;accepted.

12. Man-Son-Hing M, Laupacis A, O'Rourke K, et al. Determination of the clinical importance of study results. J Gen Intern Med. Jun 2002;17(6):469-476.

13. Peolsson A, Peolsson M. Predictive factors for long-term outcome of anterior cervical decompression and fusion: a multivariate data analysis. Eur Spine J. Mar 2008;17(3):406-414.

14. Crawford CH, 3rd, Carreon LY, Bydon M, Asher AL, Glassman SD. Impact of preoperative diagnosis on patient satisfaction following lumbar spine surgery. J Neurosurg Spine. Jun 2017;26(6):709-715.

15. Montbriand JJ, Weinrib AZ, Azam MA, et al. Smoking, pain intensity, and opioid consumption one to three months after major surgery: A retrospective study in a hospital-based Transitional Pain Service. Nicotine Tob Res. May 03 2017.

16. Bastian LA, Driscoll MA, Heapy AA, et al. Cigarette Smoking Status and Receipt of an Opioid Prescription Among Veterans of Recent Wars. Pain Med. Jun 01 2017;18(6):1089-1097.

17. Jazini E, Glassman SD, Bisson EF, Potts EA, Carreon LY. Do Former Smokers Exhibit a Distinct Profile Before and After Lumbar Spine Surgery? Spine (Phila Pa 1976). Jun 19 2017.

18. Kimura A, Shiraishi Y, Inoue H, Endo T, Takeshita K. Predictors of Persistent Axial Neck Pain After Cervical Laminoplasty. Spine (Phila Pa 1976). Jun 06 2017.

19. Mishriky BM, Habib AS. Nicotine for postoperative analgesia: a systematic review and meta-analysis. Anesth Analg. Aug 2014;119(2):268-275.

20. Abdu RW, Abdu WA, Pearson AM, Zhao W, Lurie JD, Weinstein JN. Reoperation for Recurrent Intervertebral Disc Herniation In The Spine Patient Outcomes Research Trial: Analysis of Rate, Risk Factors and Outcome. Spine (Phila Pa 1976). Jan 31 2017.

21. Leven D, Passias PG, Errico TJ, et al. Risk Factors for Reoperation in Patients Treated Surgically for Intervertebral Disc Herniation: A Subanalysis of Eight-Year SPORT Data. J Bone Joint Surg Am. Aug 19 2015;97(16):1316-1325.

22. Moon MS, Yoon MG, Park BK, Park MS. Age-Related Incidence of Cervical Spondylosis in Residents of Jeju Island. Asian Spine J. Oct 2016;10(5):857-868.

23. Friedenberg ZB, Edeiken J, Spencer HN, Tolentino SC. Degenerative changes in the cervical spine. J Bone Joint Surg Am. Jan 1959;41-A(1):61-70 .

Chapter 6

Symptomatic adjacent segment disease after anterior cervical discectomy for one-level degenerative disc disease.

Roland Donk, Wim Verhagen, Allard Hosman, André Verbeek, Ronald Bartels,

See comment in PubMed Commons below:
Published Clinical Spine Surg. 2017 Jun 9. doi: 10.1097/BSD.0000000000000551

Study design

A prospective cohort of 142 patients underwent either anterior cervical discectomy alone, anterior cervical discectomy with fusion by cage standalone or anterior cervical discectomy with arthroplasty. We then followed up on their condition a mean of 9.1 ± 1.9 years (5.6–12.2 years) later.

Objective

We aimed to evaluate the annual rate of clinically symptomatic adjacent segment disease (ASD) and to analyze predictive factors.

Summary of background data

Until recent, ASD has been predominantly evaluated radiologically. It is not known whether all patients had complaints. A frequent cited annual rate of ASD is 2.9%, but a growing number of studies report a lower annual rate. Furthermore, maintaining motion to prevent ASD is one reason for implanting a cervical disc prosthesis. However, the results of studies contradict one another.

Methods

Participants took part in a randomized controlled trial that ended prematurely because of the publication of evidence that did not justify continuation of the trial. The patients were randomly allocated to three groups, each of which received one of the abovementioned treatments. We defined symptomatic ASD as signs and symptoms caused by degeneration of an intervertebral disc adjacent to a level of previous anterior cervical disc surgery. At the last follow-up we were able to ascertain whether clinically symptomatic ASD was present in any of the participants.

Results

The overall annual rate of symptomatic ASD was 0.7%. We found no statistically significant correlations between any of the investigated factors and symptomatic ASD except for the surgical method used. Symptomatic ASD was seen less often in anterior cervical discectomy solely or anterior cervical discectomy with arthroplasty than in anterior cervical discectomy with fusion by standalone cage.

Conclusions

The annual rate of symptomatic ASD after an anterior cervical discectomy procedure was estimated to be 0.7%. This seems to be related to the procedure, although firm conclusions cannot be drawn.

Level of evidence

2 (prospective cohort)

Introduction

The ability to maintain motion and prevent adjacent segment disease (ASD) has frequently been used as an argument in favor of cervical arthroplasty and against cervical fusion. In the past, ASD was mainly a radiological diagnosis that did not take into account its clinical relevance. Radiographic evidence of degeneration does not automatically lead to clinical signs or symptoms [1]. Discussions mainly focused on the true annual incidence rate of radiologically proven ASD, which varies from 0.9% to 2.9%. With the advent of cervical arthroplasty, there was an enormous increase in interest in ASD [2]. Most previous studies have compared anterior cervical discectomy with fusion by plate fixation (ACDF) to anterior cervical discectomy with arthroplasty (ACDA). It is still unclear whether ASD is caused by ACDF or is a progression of an existing degenerative process.

Although the radiological diagnosis of ASD may be interesting, only symptomatic ASD is relevant to the patient. Since it is very difficult to attribute neck pain that requires pain relief or physiotherapy to a particular cervical segment, we define symptomatic ASD as follows: symptoms and signs in the arm that warrant intensive pain treatment or surgical treatment due to degeneration of the intervertebral disc that is adjacent to the level that has been operated on before. In this long-term follow-up study, we investigated the annual incidence of symptomatic ASD and looked for factors that are relevant to its occurrence.

Methods

The trial was approved by CMO Arnhem-Nijmegen (the regional committee overseeing research on human subjects). The study was carried out in accordance with the World Medical Association Declaration of Helsinki [3].

We designed a single center, randomized controlled trial to compare anterior cervical discectomy (ACD) alone to ACDF and ACDA [4]. Study participants were patients who presented with arm pain that lasted longer than 10 weeks, did not respond to conservative treatment and showed single-level disc degeneration and a mobile segment on dynamic lateral cervical X-rays. After patients were familiarized with the study and gave their informed consent, they were assigned randomly to either the ACD, ACDF or ACDA group. For the purpose of this study, the participants were considered as a large prospective cohort. Upright radiographs of the cervical spine were made at regular intervals and electronically stored and available using Impax ES (Agfa Web 1000 5.1, Agfa-Gevaert group, Mortsel, Belgium).

We recruited study participants from October 2003 through April 2010. After the publication of a meta-analysis of several RCTs with up to two years follow-up[5] showing that a clinical difference between ACDF and ACDA did not exist, we decided to end the trial. Because we, as investigators, were not convinced anymore of the benefit of one of the treatments compared to the others, the scientific knowledge based on which this study was designed had been altered significantly, the disc prosthesis was very costly (five times more expensive than cage standalone), and after consultation of the ethical board a justification to continue the trial according to good clinical practice could not be defined anymore. Although the inclusion was ended it was decided that the patients who were included and allocated to a treatment group would be followed till at least 5 years after the surgery. For the purpose of the current research question the sample could be considered as a cohort.

The initial surgical technique was identical for all participants. The intervertebral disc of interest was reached through a right-sided anterior approach. Radiological confirmation was obtained by placement of screw or needle into a vertebral body to avoid perforating any intervertebral disc and to prevent iatrogenic disc degeneration. A microscopic discectomy was performed, the posterior longitudinal ligament was resected and the uncinate process was reduced on both sides.

Then the intervention varied depending on the participant's group assignment. Participants in the ACD group received no further intervention. Participants in the ACDF group received a stand-alone carbon fiber reinforced polymer cage (cervical interbody (I/F) cage, DePuy Spine, Johnson and Johnson, Amersfoort, the Netherlands) that was filled with autologous cancellous bone. Participants in the ACDA group received the standard discectomy and then the implantation of a disc prosthesis (Bryan disc prosthesis, Sofamor Danek, Kerkrade, the Netherlands) following the technique described by the manufacturer.

All the participants received similar treatment after surgery. None of them were prescribed a collar and they were encouraged to be mobile as soon as possible. Participants in the ACD and ACDF groups were not allowed to take NSAIDs for three weeks postoperatively to facilitate fusion, but those in the ACDA group received a two-week prescription for meloxicam to prevent heterotopic ossification at the surgical level.

At the final follow-up, participants were asked whether they contacted a spine surgeon because of new arm pain and whether they underwent additional surgery after the index surgery. The details of those visits, radiological examinations and any extra surgical procedures were reviewed.

Apart from baseline characteristics, we collected information about the participants' history of spine (except cervical) surgery with a degenerative cause or significant low back pain. Significant low back pain was defined as low back pain that persisted for more than six weeks, necessitated conservative treatment and was memorable enough to mention when asked for a medical history.

Symptomatic ASD is defined as new pain in the arm with or without neck pain that could be attributed to abnormalities on the MRI or high-resolution CT at a level adjacent to the index level. The pain should last for more than six weeks and necessitate radiological examinations. Treatment is either conservative or surgical.

From October through December 2015, we contacted participants to ask for information about possible surgical treatment of a degenerative disc disease after the index surgery. At first this was done by sending participants a questionnaire to complete. If they did not respond, we sent a reminder. If they still did not respond, an independent secretary telephoned them and attempted to complete the questionnaire with the patient. For those patients who still could not be contacted, we obtained information about possible surgical procedures by contacting the participant's general physician.

We reviewed all preoperative MRIs and paid attention to the adjacent levels of the index surgery. Signs of degeneration were noted. We defined starting degeneration as hypo intensity of the intervertebral disc at the adjacent level compared to other levels (except for the index level) at the T2 image of the sagittal preoperative MRI. A decrease in the height of the disc level could be present. This was all in accordance with recently proposed criteria [6,7]. We addressed the shape of the cervical curvature on the preoperative cervical lateral plain X-ray in upright position and correlated the preoperative alignment of the cervical spine with the occurrence of symptomatic ASD.

An annual rate for the occurrence of ASD was calculated as follows. The number of patients suffering from ASD was divided by the total number of patients of the group of interest. This rate was divided by the total follow-up years regarding ASD, which was a summation of the maximal follow-up in years for each patient. If a patient suffered from ASD, the moment of the start of the treatment for ASD was considered as the end of the follow-up. Otherwise the moment of returning the questionnaire was the last follow-up. The final product was multiplied by 100%.

We tested the occurrence of ASD within the entire cohort of participants with each of the abovementioned items as independent variables (univariate analysis). It should be stressed that comparison of the variables between treatment groups was not done. If more than one test revealed $P < 0.1$, a multivariate analysis was initiated. If applicable, a search for confounders will be performed. Statistical analysis was performed using IBM SPSS Statistics version 22 (IBM, New York, USA). Values were represented as mean ± standard deviation (minimum-maximum) or, if not normally distributed, as median ± standard deviation (minimum-maximum). We also performed independent student-t tests or chi-square tests. If $P < 0.05$, statistical significance was assumed.

We want to emphasize that no power analysis was done for this study. However, we did perform a sample size calculation for the original randomized clinical trial based on a primary clinical outcome (i.e., a difference in Neck Disability Index). This calculation cannot be attributed to this study.

Results

Of the 142 study participants, 45 were assigned to the ACD group, 47 to the ACDF group and 50 to the ACDA group. Sex was evenly divided within the cohort was evenly split by sex: it included 71 men and 71 women. The mean age of the total population was 44.3 ± 7.0 years (18.3–59.8 years). The ages of male and female patients did not differ (P=0.236).

Mean follow-up was 9.1 ± 1.9 years (5.6–12.2 years). One patient died during follow-up due to causes unrelated to the subject of this study. We were able to retrieve data about any extra surgery for symptomatic ASD from all the patients (100%). Only two patients could not be contacted directly (one had died and one did not answer the phone); their data were collected from their general physicians.

Eight patients fulfilled the criteria for ASD and underwent surgery for a symptomatic adjacent disc disease, of whom five were allocated to the ACDF group and three to the ACD group. None of the patients who had been allocated to the ACDA group developed symptomatic ASD.

In this sample, none of the patients who developed signs and symptoms of ASD refused surgical therapy or were treated conservatively. However, three patients were operated on because of recurrent signs and symptoms at the same level: one had been allocated to the ACDA group, one to the ACD group and the other to the ACDF group. They had recurrent compression due to arthritic changes and underwent dorsal cervical foraminotomy. None of them developed symptomatic ASD.

The mean follow-up time before surgery for symptomatic ASD or final follow-up if surgery for ASD was not performed was 8.7 ± 2.3 years (0.5–12.1 years). The median time from index surgery to surgery for symptomatic ASD was 4.1 ± 2.4 years (0.5–7.2 years). The annual rate of clinical symptomatic ASD can be calculated as ((8/ 142) / 8.7) x 100%, which equals 0.7% annually.

A univariate analysis of symptomatic ASD with sex did not reveal statistical significance (P=.467); nor did age at the index surgery (P=.844), sagittal alignment of the cervical spine (P=.561), previous history of significant low back pain or spine surgery for degenerative etiology (P=.529) or preoperative MRI findings (P=.980). The operative procedure performed was statistically significantly correlated with the occurrence of symptomatic ASD (P=.043). Since we could not theoretically define any confounding factors we did not search for it statistically. The annual incidence of symptomatic ASD differed by the technique used: 0.8% for ACD, 1.2% for ACDF and 0% for ACDA.

Discussion

This study is unique because of its prospective nature, the very long follow-up time (up to more than 12 years), the extremely high follow-up rate at last follow-up, and the comparison of ACDA with ACDF with a stand-alone cage and with ACD. Until now, ACDA has always been compared to ACDF using plate fixation.

Very few participants developed signs and/or symptoms due to degeneration of an intervertebral disc adjacent to the index level. The annual rate of newly diagnosed symptomatic ASD after an anterior discectomy procedure was only 0.6%. This finding was similar to that of a retrospective study in which an annual rate of 0.32% was calculated [2]. It is also much lower than that in a study described by Hilibrand et al. [8], who reported an annual rate of 2.9%; however, the patients in that study had higher levels of pre-existing adjacent level degeneration than our cohort. They even recommended including these levels at the first surgery.

Since the diagnosis of disc degeneration was made on plain radiographs, the diagnosis of degeneration should be obvious with current MRI techniques that can show more subtle changes that indicate disc degeneration at that level. Since we were interested in newly developed degeneration at the level adjacent to the index level, patients with clear preoperative findings at the MRI indicating disc degeneration at the adjacent level were not included in the randomized controlled trial and therefore not in this cohort. Furthermore, many patients in the study by Hilibrand et al. were lost to follow-up. Calculating the annual rate of ASD in the best-case scenario (in which none of the missing patients developed ASD), the rate would be 0.4% annually after ACDF [2], which is similar to ours.

We think the overall annual rate in our study was lower than that in the study of Hilibrand et al. because of our strict inclusion criteria. We excluded patients with initially obvious radiological findings of disc protrusion or even herniation at an adjacent level. Those patients were instead offered a two-level surgical procedure, as Hilibrand et al. and others have recommended [8].

The use of a cage stand-alone for ACDF could also have contributed to a lower rate. Ahn et al. [9] recently showed that ACDF with a stand-alone cage is less likely to result in ASD than ACDF with plate fixation. They hypothesized that the stress of the adjacent segment was immediately higher after plate fixation.

Our results were based on the outcomes of participants who were included in a randomized controlled trial that ended before reaching the number of patients that should have been included based on an estimated difference in the Neck Disability Index as a primary outcome measure. The trial ended prematurely because the results of our meta-analysis did not support any advantage of ACDA over ACDF [5]. Based on a study providing level 1 evidence, we could not justify continuing the trial. The cost of the disc prosthesis, which was five times higher than the stand-alone cage, also facilitated the decision.

Several meta-analyses have demonstrated statistically significant differences in clinical outcome measurements in favor of ACDA [10-16]. In our opinion, a discussion about interpreting the results of these meta-analyses about statistically significant results without mentioning their clinical relevance is beyond the scope of this article. However, one of the meta-analyses explicitly concluded that few studies focused on ASD, and the authors recommended more research in this direction [13]. This recommendation was confirmed by the conclusion of another meta-analysis, which investigated the rate of ASD after ACDF and ACDA in single-level disease and found no difference. The authors of that study also recommended better follow-up of the participants in future studies [17].

In our study, the rate of symptomatic ASD depended on the technique. It was highest for ACDF and the annual rates for ACDA and ACD were nearly comparable. Although we found a 0.4% difference in annual rate between ACDF and ACD, we do not think it is relevant (mainly due to the sample size). It is tempting to advocate for the use of ACDA or ACD in symptomatic, single-level degenerative cervical disc disease. However, we do not feel confident enough to make such a strong recommendation due to the relatively small number of participants who underwent each procedure in our total cohort. We do find it remarkable that minor alterations at MRI were not correlated with symptomatic ASD. Either those alterations at MRI were not predictive of future degeneration or we missed them because we did not perform an MRI optimized to detect them.

Previous studies have found a correlation between cervical degenerative disc disease and lumbar degenerative disease [18-20]. However, we were unable to show a significant relationship between symptomatic ASD and a history of significant low back pain or lumbar spine surgery (17% of the patients in this cohort). It is tempting to conclude that the normal degenerative process was not relevant in comparison to the performed surgical procedure but, due to the small number of participants with symptomatic ASD, it is not appropriate to draw strong conclusions.

Our study also had a number of constraints. First, our assumption that alterations at MRI indicated the beginning of degeneration could be defined as a flaw. However, these were also criteria in proposed classifications for cervical disc degeneration at MRI [6,7].

The second constraint is our very strict definition of symptomatic ASD, which did not include patients who suffered briefly from radicular pain or neck pain due to ASD. Therefore, the annual rate may be too low because the complaints were self-limiting.

Third, our inclusion of patients who underwent reoperation at the same level is debatable. It might confound the result by increasing the chance of developing symptomatic ASD. However, the distribution of participants among groups was equal and none of them developed ASD. Therefore, we are confident that it did not act as a confounder.

Finally, this study cannot be used to advocate a clear preference for one of the methods. The inclusion of an appropriate cost analysis is also mandatory.

In conclusion, the annual rate of symptomatic ASD after an ACD procedure was estimated to be 0.7% in this prospective cohort with up to 12 years of follow-up and a 100% adherence rate. A relationship exists between the performed procedures (patients who underwent ACDF had the highest rate of symptomatic ASD and those who underwent ACDA and ACD had the lowest), although the actual differences are very small.

Table 1: *The number of patients, occurrence of ASD, annual rate of ASD and reoperations in relation to treatment modality.*

	ACDA[2]	ACDF[3]	ACD[4]	Total	Average
Number of patients	50	47	45	142	-
Symptomatic ASD[1]	0	5	3	8	-
Annual rate of ASD	0%	1.2%	0.8%	-	0.7%
Reoperation at the same level	1	1	1	3	-

[1] ASD = adjacent segment disease
[2] ACDA = anterior cervical discectomy with arthroplasty
[3] ACDF = anterior cervical discectomy with fusion by standalone cage
[4] ACD = anterior cervical discectomy;

References

1. Heo Y, Park JH, Seong HY, et al. Symptomatic adjacent segment degeneration at the L3-4 level after fusion surgery at the L4-5 level: evaluation of the risk factors and 10-year incidence. *Eur Spine J.* Nov 2015;24(11):2474-2480.

2. Donk RB, R.H.M.A. Adjacent disc degeneration in the cervical spine: personal data and a critical reappraisal of the literature. *The Internet Journal of Spine Surgery.* 2012;6(2):1-8.

3. Fuson RL, Sherman M, Van Vleet J, Wendt T. The conduct of orthopaedic clinical trials. *J Bone Joint Surg Am.* Jul 1997;79(7):1089-1098.

4. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord.* 2006;7:85.

5. Bartels RH, Donk R, Verbeek AL. No justification for cervical disk prostheses in clinical practice: a meta-analysis of randomized controlled trials. *Neurosurgery.* Jun 2010;66(6):1153-1160; discussion 1160.

6. Jacobs LJ, Chen AF, Kang JD, Lee JY. Reliable Magnetic Resonance Imaging Based Grading System for Cervical Intervertebral Disc Degeneration. *Asian Spine J.* Feb 2016;10(1):70-74.

7. Wierzbicki V, Pesce A, Marrocco L, Piccione E, Colonnese C, Caruso R. How old is your cervical spine? Cervical spine biological age: a new evaluation scale. *Eur Spine J.* Dec 2015;24(12):2763-2770.

8. Hilibrand AS, Carlson GD, Palumbo MA, Jones PK, Bohlman HH. Radiculopathy and myelopathy at segments adjacent to the site of a previous anterior cervical arthrodesis. *J Bone Joint Surg Am.* Apr 1999;81(4):519-528.

9. Ahn SS, Paik HK, Chin DK, Kim SH, Kim DW, Ku MG. The Fate of Adjacent Segments After Anterior Cervical Discectomy and Fusion: The Influence of an Anterior Plate System. *World Neurosurg.* May 2016;89:42-50.

10. Wu AM, Xu H, Mullinix KP, et al. Minimum 4-year outcomes of cervical total disc arthroplasty versus fusion: a meta-analysis based on prospective randomized controlled trials. *Medicine (Baltimore).* Apr 2015;94(15):e665.

11. Zhang Y, Liang C, Tao Y, et al. Cervical total disc replacement is superior to anterior cervical decompression and fusion: a meta-analysis of prospective randomized controlled trials. *PLoS One.* 2015;10(3):e0117826.

12. Rao MJ, Nie SP, Xiao BW, Zhang GH, Gan XR, Cao SS. Cervical disc arthroplasty versus anterior cervical discectomy and fusion for treatment of symptomatic cervical disc disease: a meta-analysis of randomized controlled trials. *Arch Orthop Trauma Surg.* Jan 2015;135(1):19-28.

13. Ren C, Song Y, Xue Y, Yang X. Mid- to long-term outcomes after cervical disc arthroplasty compared with anterior discectomy and fusion: a systematic review and meta-analysis of randomized controlled trials. *Eur Spine J.* May 2014;23(5):1115-1123.

14. Yin S, Yu X, Zhou S, Yin Z, Qiu Y. Is cervical disc arthroplasty superior to fusion for treatment of symptomatic cervical disc disease? A meta-analysis. *Clin Orthop Relat Res.* Jun 2013;471(6):1904-1919.

15. Xing D, Ma XL, Ma JX, Wang J, Ma T, Chen Y. A meta-analysis of cervical arthroplasty compared to anterior cervical discectomy and fusion for single-level cervical disc disease. *J Clin Neurosci.* Jul 2013;20(7):970-978.

16. Boselie TF, Willems PC, van Mameren H, de Bie RA, Benzel EC, van Santbrink H. Arthroplasty versus fusion in single-level cervical degenerative disc disease: a Cochrane review. *Spine (Phila Pa 1976).* Aug 1 2013;38(17):E1096-1107.

17. Verma K, Gandhi SD, Maltenfort M, et al. Rate of adjacent segment disease in cervical disc arthroplasty versus single-level fusion: meta-analysis of prospective studies. *Spine (Phila Pa 1976).* Dec 15 2013;38(26):2253-2257.

18. Kim SJ, Lee TH, Yi S. Prevalence of disc degeneration in asymptomatic korean subjects. Part 3 : cervical and lumbar relationship. *J Korean Neurosurg Soc.* Mar 2013;53(3):167-173.

19. Matsumoto M, Okada E, Toyama Y, Fujiwara H, Momoshima S, Takahata T. Tandem age-related lumbar and cervical intervertebral disc changes in asymptomatic subjects. *Eur Spine J.* Apr 2013;22(4):708-713.

20. Jacobs B, Ghelman B, Marchisello P. Coexistence of cervical and lumbar disc disease. *Spine (Phila Pa 1976).* Dec 1990;15(12):1261-1264.

Chapter 7

An assessment of the most reliable method to estimate the sagittal alignment of the cervical spine: analysis of a prospective cohort of 138 cases.

Roland Donk, Michael Fehlings, Wim Verhagen, Hisse Arnts, Hans Groenewoud, André Verbeek, Ronald HMA Bartels

Published J Neurosurg Spine. 2017 May;26(5):572-576. doi: 10.3171/2016.10.SPINE16632.
Epub 2017 Mar 3

Objective

Background: While there is increasing recognition of the importance of cervical spinal sagittal balance, there is a lack of consensus as to the optimal method to accurately assess the cervical sagittal alignment. Cervical alignment is important for surgical decision-making. Sagittal balance of the cervical spine is generally assessed using one of two methods; namely, measuring the angle between C2 and C7, and drawing a line between C2 and C7. Here, the best method to assess sagittal alignment of the cervical spine is investigated.

Methods

Data from 138 patients enrolled in a prospective study were analyzed. Two investigators independently measured the angle between C2 and C7 by Harrison's posterior tangent method, and also estimated the shape of the sagittal curve using a modified Toyama method. The mean angles of each quantitative assessment of the sagittal alignment were calculated and the results were compared. The inter-rater reliability for both methods was estimated using Cronbach's alpha.

Results

For both methods the inter-rater reliability was high: for the posterior tangent method - 0.907 and for the Toyama modified technique -0.984. For a lordotic cervical spine, defined by the modified Toyama method, the mean angle (defined by Harrison's posterior tangent method) was 23.4 ± 9.9 degrees (range: 21.7-25.2), for a kyphotic cervical spine -0.7 ± 10.7 degrees (range: -27 – 27), and for a straight cervical spine 10.5 ± 8.2 degrees (range: -11-36).

Conclusions

An absolute measurement of the angle between C2 and C7 does not unequivocally define the sagittal cervical alignment. As can been seen from the minimum and maximum values even a positive angle between C2 and C7 could be present in a kyphotic spine. For this purpose, the modified Toyama method (drawing a line from posterior inferior part of the vertebral body of C2 to the posterior upper part of the vertebral body of C7 without any measurements) is a better tool for a global assessment of cervical sagittal alignment. alignment.

Introduction

Sagittal alignment of the cervical spine is a very important consideration when undertaking surgical interventions[1]. It is one of the critical factors that influence the decision for either an anterior or a posterior approach. Furthermore, lordosis is the natural shape and restoration of the lordosis is a goal of most surgeries. Indeed, increasing evidence suggests that neurological outcomes, quality of life and the rate of adjacent segment degeneration are optimized with establishment of cervical lordosis[1,2].

Recently, more attention has been paid to global spinal balance[3,4]. Parameters like T-1 slope and sagittal vertebral axis has been introduced. These and other parameters are quantified in angles and distance (millimeters)[5-7]. To optimize and evaluate surgical treatment of cervical pathology requires not only a quantification of different parameters a, but also an assessment of the sagittal cervical alignment.

Although its importance is obvious, little attention has been given to the best way to estimate the sagittal shape of the cervical spine. Generally two methods exist: measuring the angle between C2 and C7, and drawing a line between C2 and C7[8,9]. For the measurement of the angle two methods are generally used. One is the Cobb angle method and the other the posterior tangent method. Harrison et al compared these two techniques and concluded that while both methods were reliable, the posterior tangent method better defines the changes in angular alignment between C2 and C7[10]. However, the concern remains that a given value of the angle measured by the posterior tangent method does not always describe the shape of the cervical spine correctly. A comparison between a qualitative assessment of the sagittal alignment of the cervical spine with the actual angle measured by the posterior tangent method has never been performed. A clear definition of the shape of the cervical spine is of interest for clinical use, but also for research purposes. In this study, we compared, the quantification of the angle between C2 and C7 using Harrison's posterior tangent method with a qualitative assessment of the curve using the modified Toyama method.

Methods

The study was approved by the ethical board CMO Arnhem-Nijmegen (CMO-nr: 2002/200), and has been carried out in accordance with the World Medical Association Declaration of Helsinki. Informed consent was obtained from all patients. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement has been followed[11].

The radiographs form a prospective cohort of 138 patients, who participated in a randomized controlled trial (Procon)[12] were used. This trial was registered with the

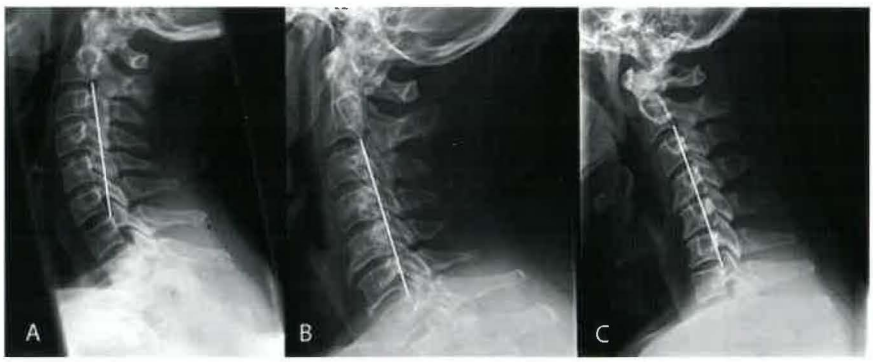
Current Controlled Trials database (https://www.isrctn.com), and its registration no. is ISRCTN41681847.

In this trial several techniques were compared: anterior cervical discectomy without fusion, anterior cervical discectomy with cage stand-alone, and anterior cervical discectomy with arthroplasty.

The preoperative upright lateral cervical radiographs and those of the last follow up moment were reviewed for this study. Radiographs were digitized and available using Impax ES (Agfa Web 1000 5.1, Agfa-Gevaert group, Mortsel, Belgium).

The curvature of the cervical spine was estimated according to a slight modification of the method by Toyama et al[8,9]. A line was drawn from the posterior and inferior part of the vertebral body of C2 to the upper posterior part of the vertebral body of C7. A lordotic curve was present if the posterior wall of the vertebral bodies of C3 to C6 were anterior of this line. The cervical spine was considered straight if the posterior part of the vertebral bodies C3 to C6 was on that line, and it was kyphotic if the posterior part of the vertebral bodies were posteriorly projected to this line (Figure 1).

Fig 1. *Radiographs showing a modification of the method proposed by Toyama et al. A cervical lordosis (A) is represented by the location of the vertebral bodies being anterior to the line drawn from C-2 to C-7; a straight spine (B) in which the posterior parts of the vertebral bodies are on that line; and a kyphotic spine (C) with the posterior parts of the vertebral bodies projecting posteriorly from that line.*



A sigmoid curve was present if the posterior part of the vertebral bodies were not on that line, but if a combination existed of posterior border of the vertebral bodies in front of the line and some behind that line.

Two investigators experienced in measuring angles (HA, RB), measured independently the angle between C2 and C7, and they also estimated the shape of the curve. The Harrison posterior tangent method was used to calculate the angle between C2 and C7 (Figure 2 and 3)[10,13].

Fig 2. *Fig 2 Radiograph showing Harrison's posterior tangent method, in which the angle between C-2 and C-7 is estimated by measuring the angle between the posterior bodies of C-2 and C-7. The lines represent the posterior bodies of C-2 and C-7. In this case the angle was 48.1° according to the Agfa Impax software.*

Fig 3. *Lateral upright radiograph of the cervical spine depicting a kyphotic spine measured with the modified Toyama approach, but with a positive angle measured by Harrison's posterior tangent method (15.8°)*

The mean angle derived from Harrison's method was then calculated for lordotic, straight and kyphotic cervical spines (as assessed by the modified Toyama method).

For statistical analysis IBM SPSS Statistics version 22 (IBM, New York, USA) was used. Values were represented as mean ± standard deviation (minimum-maximum). For inter-rater reliability Cronbach's alpha was estimated for the measurements at the preoperative lateral radiographs.

Results

The mean age of the entire group at surgery was 44.3 ± 7.0 years (18.3-59.8), and the female to male ratio 1:1. A total of 276 lateral radiographs of the cervical spine were retrieved and used. The mean time between the first radiograph, and radiograph at the last follow up was 25.4 ± 18.4 months.

Using the modified Toyoma method lordosis was found at 128 radiographs, kyphosis at 27, and a straight spine at 121 radiological exams (Table 1).

Table 1: *Values of the angle between C2 and C7 as measured by the posterior tangent method in relation to categorization of lordotic, straight or kyphotic sagittal alignment of the cervical spine.*

	Lordosis	Straight	Kyphosis
Angle C2-C7 posterior tangent method			
Mean	23.4	10.5	-2.2
Standard deviation	9.9	8.2	9.2
Minimum	0.4	-10.6	-16.1
Maximum	52.4	35.9	16.9
Total number of investigations	128	121	27

A sigmoid curve was not found. Comparing the pre – and postoperative radiographs, none of the lordotic or straight cervical spines became kyphotic. Half of the kyphotic spines remained kyphotic whereas the remaining was either straight or lordotic.

For the lordosis the mean angle was 23.4 ± 9.9 degrees (0.4 - 52.4), for kyphosis -2.2 ± 9.2 degrees (-16.1 - 16.9), and for a straight cervical spine 10.5 ± 8.2 degrees (-10.6 -35.9). An example of a kyphotic cervical spine and a positive angle between C2 and C7 is shown in Figure 3. The Cronbach's Alpha was 0.907 for the measurements for the angle C2-C7, and for the modified Toyama method 0.984.

Based on these data, we calculated the sample size needed and also the power based on the current sample size according to the method proposed by Bonnet and Wright[14]. Considering the modified Toyama method resulting in one question with three options (lordotic, straight, kyphotic) and assuming an expected intra-rater reliability of at least 0.8, $\alpha = 0.05$, and a power of 80% the sample size needed would be 21. Of note, the power of the results in a group of 276 is 100%.

Discussion

This study clearly disclosed the discrepancy between the absolute values of the angle between C2 and C7 and the qualification of the cervical sagittal alignment expressed as lordotic, straight, or kyphotic. The modified Toyama method has been shown to be reliable and easy to assess. This is in daily practice and for the individual patient of utmost importance to offer the best approach if surgery is needed.

Recently, global spinal alignment is subject of many investigations. As a part of these investigations attention has been paid to quantify cervical alignment by using parameters as cervical sagittal vertical axis (SVA), cervical lordosis, and T1 slope. A correlation has been found between these parameters and health-related quality of life[15,16]. The effect of surgical correction of thoracic or lumbar deformities is well known and even predictive models have been developed[17-19]. In these studies the assessment of the alignment of the cervical spine was never addressed. A possible explanation is that in case of deformity surgery one of the goals is to maintain horizontal gaze while having the possibility to flex in the cervicothoracic region to allow walking without stumbling while ambulating[4]B. G.</author><author>Challier, V.</author><author>Henry, J. K.</author><author>Oren, J. H.</author><author>Spiegel, M. A.</author><author>Vira, S.</author><author>Tanzi, E. M.</author><author>Liabaud, B.</author><author>Lafage, R.</author><author>Protopsaltis, T. S.</author><author>Errico, T. J.</author><author>Schwab, F. J.</author><author>Lafage, V.</author></authors></contributors><auth-address>*Spine Service, Hospital for Special Surgery, New York, NY, United States daggerSpine Unit 1, Orthopedic Surgery Department, Bordeaux University Hospital, Bordeaux, France double daggerDepartment of Orthopaedic Surgery, NYU Langone Medical Center, New York, NY, United States section sign-Memorial Sloan Kettering Cancer Center, New York, NY.</auth-address><titles><title>Predicting Cervical Alignment Required to Maintain Horizontal Gaze Based on Global Spinal Alignment</title><secondary-title>Spine (Phila Pa 1976. The shape of the cervical spine is in these cases less important than the other quantitative measurements. Furthermore, the regional curves have a wide variation in undulating while the measurements for global sagittal balance are maintained in narrower range[16].

We fully endorse the importance of global spinal alignment. However, a major shortcoming of the quantifications mentioned above is that the actual alignment is not assessed. For example, a C2-C7 SVA of 15 millimeters does not automatically mean a kyphotic spine or otherwise. For most prevalent diseases like cervical spondylotic myelopathy the sagittal alignment is one of the factors that determines the optimal surgical approach. Sagittal alignment is of utmost importance for surgical planning and to guide decision-making as to whether to offer the patient an anterior, a posterior, or a combined approach.

It was clearly shown that the absolute angle between C2 and C7 using a technique that accurately depicts the cervical angle, locally and overall[10] did not reflect the alignment of the cervical spine. The angle of the vertebral body of C2 and C7 can nearly be zero or even positive indicating a straight of even a lordotic spine, whereas the vertebral bodies in between (C3-C6) contributed to a true cervical kyphosis.

For scientific but also practical purposes a clear and reproducible method for estimation of the cervical sagittal alignment is warranted. Toyama described a method drawing a line from the posterior inferior part of the vertebral body of C2 to the upper posterior edge of the vertebral body of C7. Then distances between the line and the posterior border of the vertebral body should be measured, and if a certain distance is (more than 2 mm) reached a lordotic, kyphotic or sigmoid curve is present[9]. Measuring these distances is prone to error and is not very convenient even in the time of digitization.

Benzel described also a method in which a line was drawn between the posterior inferior part of the vertebral body of C2 to the lower posterior edge of the vertebral body of C7[20]. This can be troublesome since C7 is very often not completely shown at a lateral X Ray. We modified the method by Toyama by using the same line without the measurements[8]. If the posterior wall of the vertebral bodies C3 through C6 were in front of that line cervical lordosis existed, were they on that line it was a straight spine and were they behind it we called it kyphosis.

Another method is estimating the angle in the sagittal plane between C2 and C7. Compared to the Cobb method the posterior tangent method was more accurate[10]. It is a general belief that a positive angle corresponds with a lordotic curve. The results of this study counter this belief.

This can be explained by the fact that the measurement between the C2 and C7 is reflecting the angle between these two vertebral bodies. The position of these bodies is more or less independent of that of the intermediate vertebral bodies. For example the posterior walls of C3 to C7 can be exactly co-linear, but if C2 is angulated posteriorly the angle between C2 and C7 can be very positive whereas the line drawn by the modified Toyama method projects exactly over the posterior wall of the intermediate vertebral bodies. In this way, a positive angle in the posterior tangent method is combined with a straight cervical spine. This holds also true for a kyphotic position of the vertebral bodies of C3, C4, C5 and C6 with compensation at C2c3 and c6c7. Locally at C2C3 and C6C7 will be very positive, but between C2 and C7 these can still be positive suggesting lordosis whereas using the modified Toyama method an overall kyphotic curve is present. In other words, the posterior tangent method derives all its information locally at the end points and infers a conclusion about the entire cervical region while the modified Toyama is obtains information from each segment.

Considering the high inter-rater reliability, the posterior tangent method as well as the modified Toyama method was very useful. However, the description of the curvature of the cervical spine is more reliable represented by the modified Toyama method. With the posterior tangent method the curvature between C2 and C7 was neither accurately described, e.g. in a kyphotic curve assessed by the modified Toyama method still a positive angle could be measured between C2 and C7.

This has practical implications. In each patient this convenient and reproducible method will facilitate the decision whether to perform surgery anteriorly or posteriorly. Also in research the description of the curve through this method is more reliable than the absolute value of the angle. Since this method only assesses the shape of the cervical spine without quantifica-

tion, future investigations will be necessary that will facilitate its introduction in global spinal alignment measures. The combination of an assessment of the alignment of the cervical spine with a quantitative measure would be of great benefit for the comparison of various surgical approaches, for outcome research and foe educational purposes. A suggestion could be introducing it as a modifier in combination with the C2-C7 SVA.

It could be argued that the results of this studied were biased since the patients participated to a study. This might introduce selection bias. For example, the population was relatively young although the mean age resembled that of patients suffering from symptomatic cervical degenerative disc disease. Since the cervical alignment was not an inclusion or exclusion criterion, the relatively few kyphotic patients in this series (9.7%) might raise some discussion. However, this number also represented the normal prevalence[21]. The strength of the study will not be influenced by the unevenly distributed sample size, since the sample size was much larger than needed. Furthermore, the investigated methods can be used in any patient with any cervical curve. The high intra-rater reliability, also in distinguishing the kyphotic and straight shape and the post hoc power analysis of 100%, provide a level of confidence in our conclusions.

Although the method is really simple to assess and has a high intra-rater reliability, we acknowledge that widespread application will reduce the intra-rater reliability.

Conclusions

For practical purposes the modified Toyama method by drawing a line from posterior inferior part of the vertebral body of C2 to the posterior upper part of the vertebral body of C7 without any measurements is accurate, straightforward, and reliable to qualify the sagittal alignment of the cervical spine. This technique can be readily and easily incorporated into preoperative surgical planning.

Reference

1. Shamji MF, Mohanty C, Massicotte EM, Fehlings MG. The Association of Cervical Spine Alignment with Neurologic Recovery in a Prospective Cohort of Patients with Surgical Myelopathy: Analysis of a Series of 124 Cases. World Neurosurg. Sep 25 2015.

2. Hansen MA, Kim HJ, Van Alstyne EM, Skelly AC, Fehlings MG. Does postsurgical cervical deformity affect the risk of cervical adjacent segment pathology? A systematic review. Spine (Phila Pa 1976). Oct 15 2012;37(22 Suppl):S75-84.

3. Diebo BG, Oren JH, Challier V, et al. Global sagittal axis: a step toward full-body assessment of sagittal plane deformity in the human body. J Neurosurg Spine. May 20 2016:1-6.

4. Diebo BG, Challier V, Henry JK, et al. Predicting Cervical Alignment Required to Maintain Horizontal Gaze Based on Global Spinal Alignment. Spine (Phila Pa 1976). May 18 2016.

5. Gillis CC, Kaszuba MC, Traynelis VC. Cervical radiographic parameters in 1- and 2-level anterior cervical discectomy and fusion. J Neurosurg Spine. May 6 2016:1-9.

6. Iyer S, Lenke LG, Nemani VM, et al. Variations in Occipitocervical and Cervicothoracic Alignment Parameters based on Age: A Prospective Study of Asymptomatic Volunteers using Full-Body Radiographs. Spine (Phila Pa 1976). Apr 22 2016.

7. Iyer S, Lenke LG, Nemani VM, et al. Variations in Sagittal Alignment Parameters based on Age: A Prospective Study of Asymptomatic Volunteers using Full-Body Radiographs. Spine (Phila Pa 1976). Apr 20 2016.

8. Bartels RH, Donk RD, Feuth T. Subsidence of stand-alone cervical carbon fiber cages. Neurosurgery. Mar 2006;58(3):502-508; discussion 502-508.

9. Toyama Y, Matsumoto M, Chiba K, et al. Realignment of postoperative cervical kyphosis in children by vertebral remodeling. Spine (Phila Pa 1976). Nov 15 1994;19(22):2565-2570.

10. Harrison DE, Harrison DD, Cailliet R, Troyanovich SJ, Janik TJ, Holland B. Cobb method or Harrison posterior tangent method: which to choose for lateral cervical radiographic analysis. Spine (Phila Pa 1976). Aug 15 2000;25(16):2072-2078.

11. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. Oct 16 2007;147(8):573-577.

12. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. BMC Musculoskelet Disord. 2006;7:85.

13. Scheer JK, Tang JA, Smith JS, et al. Cervical spine alignment, sagittal deformity, and clinical implications: a review. J Neurosurg Spine. Aug 2013;19(2):141-159.

14. Bonnet GBW, T.A. Cronbach's alpha reliability: interval estimation, hypothesis testing, and sample size planning. J Organiz Behav. 2015;36:3-15.

15. Ames CP, Blondel B, Scheer JK, et al. Cervical radiographical alignment: comprehensive assessment techniques and potential importance in cervical myelopathy. Spine (Phila Pa 1976). Oct 15 2013;38(22 Suppl 1):S149-160.

16. Kuntz Ct, Levin LS, Ondra SL, Shaffrey CI, Morgan CJ. Neutral upright sagittal spinal alignment from the occiput to the pelvis in asymptomatic adults: a review and resynthesis of the literature. J Neurosurg Spine. Feb 2007;6(2):104-112.

17. Passias PG, Oh C, Jalai CM, et al. Predictive Model for Cervical Alignment and Malalignment Following Surgical Correction of Adult Spinal Deformity. Spine (Phila Pa 1976). Apr 19 2016.

18. Protopsaltis T, Bronsard N, Soroceanu A, et al. Cervical sagittal deformity develops after PJK in adult thoracolumbar deformity correction: radiographic analysis utilizing a novel global sagittal angular parameter, the CTPA. Eur Spine J. Jul 20 2016.

19. Lafage R, Challier V, Liabaud B, et al. Natural Head Posture in the Setting of Sagittal Spinal Deformity: Validation of Chin-Brow Vertical Angle, Slope of Line of Sight, and McGregor's Slope With Health-Related Quality of Life. Neurosurgery. Jul 2016;79(1):108-115.

20. EC B. Biomechanics of Spine Stabilization. 2001; Rolling Meadows, IL: AANS.

21. Moon BJ, Smith JS, Ames CP, et al. Prevalence and type of cervical deformities among adults with Parkinson's disease: a cross-sectional study. J Neurosurg Spine. Apr 2016;24(4):527-534.

Chapter 8

Cervical sagittal alignment after different anterior discectomy procedures for single level cervical degenerative disc disease: randomized controlled trial

Roland Donk, Hisse Arnts, Wim Verhagen, Hans Groenewoud, André Verbeek, Ronald Bartels

Published in Acta Neurochir (2017) 159:2359–2365
DOI: 10.1007/s00701-017-3312-z

Abstract

Background

The effect on cervical sagittal alignment of anterior cervical discectomy without fusion (ACD), ACD with fusion by cage stand-alone (ACDF) or with arthroplasty (ACDA) is not known, and is subject of this study.

Methods

142 adult patients with single level cervical disc disease were at random allocated to different procedures: ACD (45), ACDF (47) and ACDA (50). Upright cervical spine radiographs were obtained. Angles of the involved angle and the angle between C2 and C7 were determined.

Result

After a mean follow-up of 25.4 ± 18.4 months, the angles of the involved level comparing ACD with ACDA and ACD with ACDF were different, reaching statistical significance. However, the angle between C2 and C7 did not differ between groups, nor between preoperative values and at follow-up.

Conclusions

Irrespective of the technique used for anterior cervical discectomy for single level degenerative disc disease, the alignment of the cervical spine is unaltered.

Introduction

Global sagittal balance of the spine is currently a main focus of research. Several studies have been published, stressing the importance of correct sagittal balance in relation to the quality of life[1-3]. As a consequence, the attention for cervical alignment is also increasing. Measurements such as the T1 slope and C2-C7 sagittal vertical axis (SVA) have been introduced. A good correlation between these measurements on full spine radiographs and the "classical" measurements on sagittal cervical radiographs have been established[4].

Anterior cervical approaches for degenerative disc disease are very familiar to spine surgeons, and might affect the cervical sagittal alignment. The first descriptions of anterior cervical discectomy without fusion (ACD) and ACD with fusion (ACDF) have been reported nearly at the same time by Hirsch[5], and Cloward[6], respectively. Local kyphosis had already been mentioned by Hirsch in his original article, as had fusion of the involved level[5]. Cloward stated that prevention of osteophytic spur formation could be prevented by fusion[6]. Kyphosis as well as prevention

of spur formation might be the reason that ACDF became more popular, and is considered the golden standard.

Although proper investigations about the clinical superiority of ACDF have never been performed, ACD has almost been abandoned in clinical practice. This is remarkable, since ACD provides a similar adequate decompression without the need for any implant[7,8].

Cervical sagittal balance has been investigated in patients after ACDA and ACDF[9-12]. However, the effect of ACDF with stand-alone cage or ACD on cervical sagittal balance has never been evaluated, nor compared with ACDA. This study fills this scientific gap in literature.

Methods

The Ethics Committee CMO Arnhem-Nijmegen approved the trial (CMO-nr, 2002/200). The study was carried out in accordance with the World Medical Association Declaration of Helsinki[13]. A single-center, randomized controlled trial was designed comparing ACD, ACDF and ACDA[14]. From October 2003 till April 2010, patients were included in the study after having signed informed consents. However, inclusion was prematurely ended after publication of a meta-analysis comparing ACDF and ACDA[15], since we could not justify continuing the trial according to standards of Good Clinical Practice.

Patients with (1) arm pain, not responding to conservative treatment and (2) that lasted longer than 10 weeks, with (3) single level disc degeneration and (4) a mobile spine on dynamic lateral cervical X-rays were included in the (PROCON)-trial. After informed consent, they were allocated to ACD, ACDF or ACDA. For randomization, a closed envelope method was used. A medical secretary unaware of the purpose of the study disclosed the decision. Because of radiological follow-up, neither patients nor investigators were blinded. However, the surgeons and investigators did not have any preference for any surgical method. Clinical and radiological follow-up was initially scheduled at regular intervals: 1 day postoperatively, 6 weeks, 3 months, 1 year and 2 years postoperatively. During the trial, the follow-up protocol was altered in consultation with the ethics committee, and after requests of several patients, who asked whether outpatient clinical visits were necessary. They preferred completing the questionnaires at home. The protocol was changed, and patients were asked to visit the outpatient clinic preferentially till 1 year postoperatively. Afterwards it was voluntary. If they decided to complete the questionnaires at home, they were sent to them by flat mail.

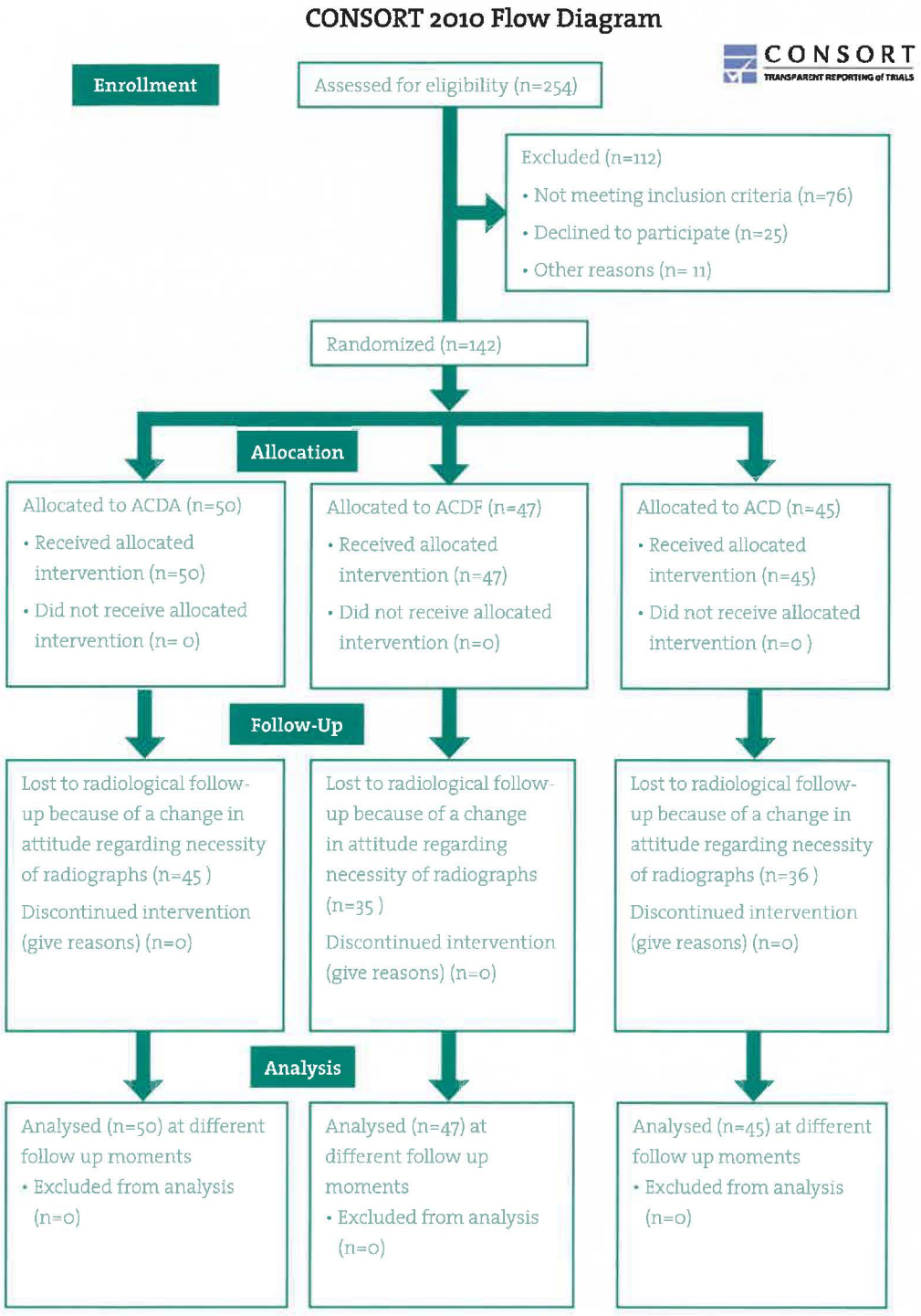
Upright cervical spine radiographs were made. Radiographs were digitized and available using Impax ES (Agfa Web 1000 5.1, Agfa-Gevaert group, Mortsel, Belgium). The Harrison posterior tangent method was used as an estimate for measuring cervical alignment[1,16]. A positive angle resembled lordosis, whereas a negative one defined kyphosis. The curvature was also estimated using a slight modification of the method by Toyama et al[17]. A line was drawn from the posterior and inferior part of the vertebral body of C2 to the upper posterior part of the vertebral body of C7[18,19]. The curvature was defined as lordotic if the posterior wall of the vertebral bodies of C3 to C6 were anterior of this line. The cervical spine was considered straight if the posterior part of the vertebral bodies C3 to C6 were on that line, and kyphotic if the posterior parts of the vertebral bodies were posterior to this line.

Two investigators, experienced in measuring spine angles (HA, RB), independently measured the angle of the involved levels as well as the angle between C2 and C7, both preoperatively and at postoperative follow-up (FU) times. They also estimated the curvature of the whole cervical spine. For statistical analyses, SAS version 9.2(SAS Institute Inc. Cary NC, USA) was used. Inter-rater reliability was assessed by calculating Cronbach's alpha for the measurements at the involved level and for the C2-C7 angle. For comparison of baseline characteristics, ANOVA or chi square tests were used. Data were represented as mean ± standard deviation (minimum-maximum). When appropriate, 95% confidence intervals were provided. Statistical significance was assumed if $P < 0.05$.

Results

Of the 142 patients who were included in the study, 45 were allocated to ACD, 47 to ACDF and 50 to ACDA. The mean age of the entire group was 44.3 ± 7.0 years (18.3-59.8), and the female-to-male ratio 1:1. Baseline characteristics regarding age or gender did not statistically differ (P= 0.287 respectively P= 0.853). The baseline characteristics for the groups are presented in Table 1. Mean radiological follow-up was 25.4 ± 18.4 months, whereas mean clinical follow-up was 9.1 ± 1.9 years (5.6 – 12.2). The flow chart according to Consort is represented in Figure 1.

Figure 1: Flow chart of included patient according to Consort 2010



Actual radiological follow up differed from the follow-up protocol. This variation was due to availability of the preferences of physicians and patients,

Among the three groups, there was no statistically significant difference in follow-up

(P = 0.18). Moreover, no preoperative statistically significant differences could be detected among treatment groups in either the mean angle of the involved (affected) level and mean global sagittal (C2-C7) alignment (Table 1).

Table 1: *Distribution of mean age, gender (percentage of column), angles and total number of patients related to procedure*

	ACDA	ACDF	ACD
Male	24 (48 %)	25 (53,2 %)	22 (48.9 %)
Female	26 (52 %)	22 (46,8 %)	23 (51.1 %)
Age ??	53.6 ± 6.9	52.2 ± 8.1	54.0 ± 6.1
Angle involved level	2.3±4.6	3.2±5.3	1.7±5.6
Angle C2-C7	14.8±13.6	12.8±12.4	16.2±12.6
Level C4C5	0	2	1
Level C5C6	21	19	26
Level C6C7	29	26	18
Total Number	50	47	45

The Cronbach's alpha was 0.837 for measurements of the involved level and 0.907 for the measurements for the C2-C7 angle, indicating a high inter-rater reliability. The mean values for the follow-up angles of the involved level and C2-C7 are presented in Table 2 and Table 3.

Table 2: *Angle in degrees at the involved level (mean ± SD) at the different follow-up (FU) times. Number of calculations (N) was also represented.*

	Time postoperatively (mean ± std dev) (weeks)	N	ACDA	ACDF	ACD	P - value
Preoperative	N/A	138	2.3±4.6	3.2±5.3	1.7±5.6	.393
Directly postoperative	N/A	140	2.5±6.1	8.4±5.2	-4.1±5.8	.000
FU2	9.3 ± 8.6	140	1.4±5.5	4.7±6.1	-2.4±5.5	.000
FU3	47.0 ± 35.2	131	1.4±5.5	4.7±6.1	-2.4±5.5	.000
FU4	134.4 ± 75.6	83	0.3±6.1	3.6±6.9	-2.9±4.1	.001
FU5	147.8 ±57.3	27	1.8±5.4	3.7±4.3	-3.6±4.2	.039

Table 3: *angle in degrees at the involved level (mean ± std dev) at the different follow up (FU) moments. Number of calculations (N) was also represented.*

	Time postoperatively (mean ± std dev) (weeks)	N	ACDA	ACDF	ACD	P - value
Preoperative	N/A	135	14.8±13.6	12.8±12.4	16.2±12.6	0.459
Directly postoperative	N/A	140	12.8±10.7	13.3±11.4	9.8±10.4	0.260
FU2	9.3 ± 8.6	139	15.0±11.6	16.8±10.3	13.0±10.5	0.265
FU3	47.0 ± 35.2	128	14.0±12.7	16.6±12.8	16.5±11.0	0.524
FU4	134.4 ± 75.6	83	14.0±12.9	18.3±13.1	17.1±11.1	0.386
FU5	147.8 ±57.3	26	14.9±11.5	16.2±13.5	17.3±6.4	0.919

The overall mean angles were 2.4 ± 5.2 degrees and 14.5 ± 12.9 degrees. One day postoperatively, a clear difference was observed in the angle of the involved level compared with the preoperative one. At the following follow-up times, a gradual decline to the baseline preoperative value was seen in all groups. In the ACD group, this was less prominent.

Figure 2: Graphs depicting the preoperative measurements of the sagittal angles at the surgical level and the values at the different follow-up times

Figure 2 clearly depicts that directly postoperative (FU1: day 1 postoperatively) a statistically significant transformation is found at the involved level in patients who underwent ACD and ACDF. This was not the case for ACDA. In the ACD group, a more negative angle was found directly postoperative, indicating the introduction of local kyphosis. However, after FU2 (9.3 ± 8.6 weeks postoperatively) changes of the angle at the involved levels did not occur anymore. Therefore, the angle measured at FU2 seemed to represent the final situation. Between FU2 and FU5 (147.8 ± 57.3 weeks postoperatively) the difference did not reach statistical significance for any of the groups. Since only 28 patients were evaluated after FU5 for reasons explained in the discussion, the 95% confidence intervals were very wide. Therefore, we compared the difference with the measurements at FU4 (134.4 ± 75.6 weeks postoperatively). The differences between FU2 and FU4 were also not statistically significant (except for the angle at the involved level for ACDA). The mean differences within each group at the successive follow-up times could be considered as measurement error (Table 4).

Table 4: *Differences in angles in degrees at the involved level (mean (95% CI)) for the different follow up (FU) times (PREOP: preoperative; POSTOP: 1 day postoperatively*

	ACDA	ACDF	ACD
POSTOP-PREOP	0.1 (-1.7 ; 2.0)	5.2 (3.8 ; 6.6)	-5.2 (-7.3 ; -3.1)
FU2-POSTOP	-1.2 (-2.2 ; -0.2)	-3.3 (-4.7 ; -1.8)	0.9 (-0.3 ; 2.1)
FU3-FU2	0.2 (-0.9 ; 1.2)	-0.2 (-1.4 ; 1.0)	0.8 (-1.2 ; 2.8)
FU4-FU3	-1.5 (-2.5 ; -0.5)	-0.7 (-1.8 ; 0.5)	-0.4 (-1.6 ; 0.8)
FU5-FU4	0.6 (-1.0 ; 2.2)	-0.7 (-2.5 ; 1.1)	-0.0 (-2.0 ; 2.0)
FU4-PREOP	-2.2 (-4.2 ; -0.2)	1.2 (-0.7 ; 3.2)	-4.1 (-6.4 ; -1.8)

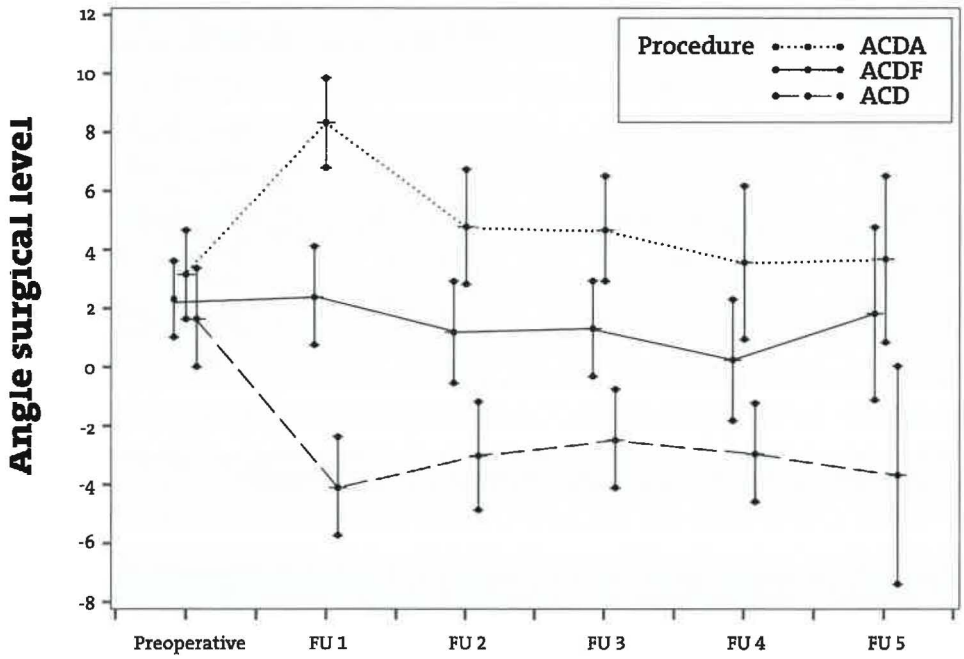
The same observation was made for the C2-C7 angle (Table 5).

Table 5: *differences in angles in degrees between C2 and C7 (mean ± std dev) for the different follow up (FU) moments (PREOP: preoperative; POSTOP: 1 day postoperatively)*

	ACDA	ACDF	ACD
POSTOP-PREOP	-1.4 (-4.8 ; 2.1)	-0.1 (-3.1 ; 3.0)	-6.0 (-9.9 ; -2.1)
FU2-POSTOP	2.2 (-0.7 ; 5.1)	3.8 (0.9 ; 6.6)	3.0 (-0.4 ; 6.3)
FU3-FU2	-0.6 (-3.3 ; 2.1)	0.2 (-2.0 ; 2.4)	3.4 (0.1 ; 6.5)
FU4-FU3	-1.3 (-4.5 ; 1.9)	1.4 (-2.6 ; 5.5)	-0.4 (-4.7 ; 3.9)
FU5-FU4	2.8 (-2.7 ; 8.3)	-4.1 (-10.8 ; 2.5)	-3.1 (-9.6 ; 3.4)
FU4-PREOP	-0.9 (-4.4 ; 2.5)	7.3 (3.7 ; 10.9)	1.0 (-4.0 ; 5.9)

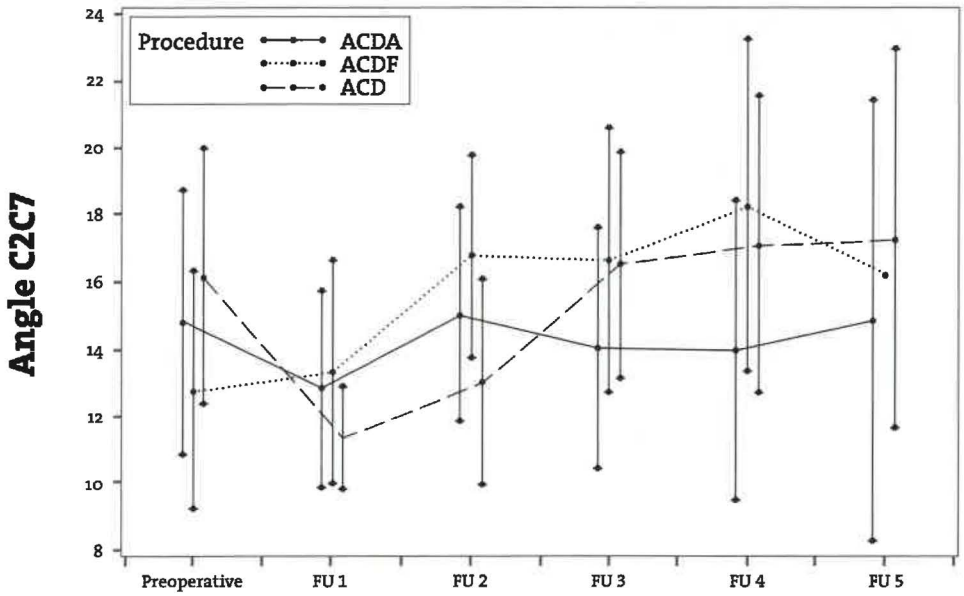
Between the treatment groups, differences existed when comparing the local angle of the involved level. This can clearly be seen in figure 2.

Figure 2: Graphs depicting the preoperative measurements of the sagittal angles at the surgical level and the values at the different follow-up times



The mean values for ACDA were 1.4 ± 5.5 degrees for ACDF 4.7 ± 6.1 and for ACD -2.4 ± 5.5 ($P < 0.0001$) at FU3 (N=131). For the C2-C7 angle statistical significance was not reached when comparing the groups ($P= 0.305$) as depicted in figure 3.

Figure 3: Graphs showing the preoperative measurements of the sagittal angles at C2-C7 and the values at the different follow-up times.



The shape of the cervical spinal curve did not change in 111 patients during follow-up. In 31 patients the shape did alter (Table 6), but none of them became kyphotic.

Table 6: *The pre- and postoperative sagittal curve of the cervical spine defined as lordotic, straight and kyphotic*

Procedure	Preoperative shape	Postoperative shape		
		Kyphosis	Straight	Lordosis
ACDA	Kyphotic	4	6	1
	Straight	0	15	4
	Lordotic	0	23	0
	Unkown	0	2	0
ACDF	Kyphotic	3	1	2
	Straight	0	12	7
	Lordotic	0	1	17
ACD	Kyphotic	4	1	0
	Straight	0	21	5
	Lordotic	0	12	0
	Unknown	0	1	0

The ultimate shape of the curvature was not dependent upon the used technique, but (statistically significant) upon the preoperative shape ($P<0.001$).

Discussion

Most radiological studies in relation to arthroplasty focus on ROM and movement of the adjacent levels in comparison to ACDF with plate fixation. A few studies described the sagittal cervical balance after arthroplasty in comparison ACDF with plate fixation[11, 20, 21]. Retrospective cohorts have also been published [9, 22, 23]. A recent systematic review showed that after ACDA the alignment of the cervical spine tended to become kyphotic[24], which concurred with our results.

This is the first study that evaluated the effect of ACD, ACDF with a stand-alone cage and ACDA on cervical sagittal alignment at both the involved (affected) level and the cervical spine (C2-C7). While ACD is a well-known procedure, it has received little attention in the last ten years. The research focused on comparing ACDF with plate fixation with ACDA. This study is unique because of its prospective nature and the comparison of ACDA with ACDF with a stand-alone and with ACD.

Two remarkable findings of this study should be mentioned. First, though the angle at the involved level became more lordotic after ACDF and more kyphotic after ACD, it tended to normalize to its former preoperative value at approximately 9-week's evaluation for ACDF. For ACD the change was only minimal (one degree) and remained locally kyphotic. Moreover, after nearly 1 year, no differences between the three procedures could be detected. Second, global cervical lordosis was not affected by the three different techniques. These findings can be explained by the natural mechanism of the human body to maintain the head in neutral axis in the horizontal

plane optimal for the visiovestibular system and restore sagittal balance. To maintain the global sagittal balance, it seems logical that after a relatively small disturbance at the involved level, it will be locally resolved and affect the whole spine. It should be emphasized that the current investigation has only focused on radiological cervical sagittal balance and the effect of time on both the involved (affected) levels and the global cervical spinal curvature. However, Carreon et al. showed a good correlation of the measurements on the lateral full spine radiographs compared to the dedicated lateral cervical radiographs[4].

Statements about clinical performance or quality of life after the different procedures cannot be made from these results. However, our study clearly shows that the way of performing an anterior cervical discectomy is not affecting cervical lordosis in time. Therefore, one might assume that any eventual difference in clinical result should not be attributed to sagittal alignment. Considering the goal of our study, the lack of correlation with clinical outcome is not a weakness of the study but a strength, since the focus on the sagittal balance contributed to a clear interpretation of the results and discussion.

The chosen procedure only seemed to affect the angle of the involved segment to a minor aspect. Therefore, the argument of restoring lordosis by increasing it at one involved level for one-level degenerative disease is at least debatable. A limitation of our study is the loss of patients for radiological follow-up resulting in larger standard deviations. However, we feel confident that the findings represent the actual situation since the angle at the surgical level did not change significantly after the second follow-up time and the sagittal angle of C2-C7 at the last follow-up was similar to the preoperative one.

Furthermore, while the study was ongoing, more patients questioned why they should visit the hospital since they could report their outcome measurements at home. Many of them found the radiographic control exams irrelevant unless symptoms occurred. Therefore, we adapted the protocol. In the literature arguments were found to support this alteration [25]. Furthermore, the primary outcome measure of the trial was clinical outcome. To optimize the participation of the patients, we were willing to facilitate their cooperation.

Not including the C2-C7 SVA, T1 slope or other measurements in our study might be debated. However, since a good correlation was estimated with Cobb angles[4, 26], and we were interested also in the angle of the involved angle, we decided to measure the C2-C7 angle. Furthermore, attention was given to changes within groups of patients. Due to the kind of procedure, we did not assume that a clinically significant translation within the cervical spine would occur since posterior elements remained untouched and the disruption of the anterior part was minimal. Otherwise, measurement of the C2-C7 SVA would be more appropriate[27, 28].

Our study shows that the most important changes to cervical alignment took place 1 day after the surgery and in the immediate weeks thereafter. From FU2 (approximately 9 weeks), changes did not occur anymore in local angle or in the global cervical sagittal alignment. The major strength of the study is the design, facilitating the formation of three groups, with comparable

radiological baseline characteristics. This made the study unique. The high inter-rater reliability also contributed to the strength of this study.

In conclusion, at longer follow-up sagittal cervical alignment was not affected by the procedure for cervical anterior discectomy. Despite the initial increase or decrease of lordosis at the involved level, the tendency developed to restore local cervical alignment to the preoperative situation. This could be interpreted as a natural inborn mechanism to restore a longstanding situation to which the body has been accustomed. Restoring local cervical lordosis as an argument to promote a certain procedure for a single level cervical degenerative disc disease is at least debatable.

References

1. Scheer JK, Tang JA, Smith JS, et al. Cervical spine alignment, sagittal deformity, and clinical implications: a review. *J Neurosurg Spine.* Aug 2013;19(2):141-159.

2. Roguski M, Benzel EC, Curran JN, et al. Postoperative cervical sagittal imbalance negatively affects outcomes after surgery for cervical spondylotic myelopathy. *Spine (Phila Pa 1976).* Dec 1 2014;39(25):2070-2077.

3. Harroud A, Labelle H, Joncas J, Mac-Thiong JM. Global sagittal alignment and health-related quality of life in lumbosacral spondylolisthesis. *Eur Spine J.* Apr 2013;22(4):849-856.

4. Carreon LYS, C.L.; Dimar II, J.R.; Glassman, S.D. Correlation of cervical sagittal alignment parameters on full-length spine radiographs compared with dedicated cervical radiographs. *Scoliosis Spinal Disord.* 2016:11-12.

5. Hirsch C, Wickbom I, Lidstroem A, Rosengren K. Cervical-Disc Resection. A Follow-up of Myelographic and Surgical Procedure. *J Bone Joint Surg Am.* Dec 1964;46:1811-1821.

6. Cloward RB. The anterior approach for removal of ruptured cervical disks. *J Neurosurg.* Nov 1958;15(6):602-617.

7. Robertson JT, Johnson SD. Anterior cervical discectomy without fusion: long-term results. *Clin Neurosurg.* 1980;27:440-449.

8. Maurice-Williams RS, Dorward NL. Extended anterior cervical discectomy without fusion: a simple and sufficient operation for most cases of cervical degenerative disease. *Br J Neurosurg.* Jun 1996;10(3):261-266.

9. Guerin P, Obeid I, Gille O, et al. Sagittal alignment after single cervical disc arthroplasty. *J Spinal Disord Tech.* Feb 2012;25(1):10-16.

10. Kim SW, Shin JH, Arbatin JJ, Park MS, Chung YK, McAfee PC. Effects of a cervical disc prosthesis on maintaining sagittal alignment of the functional spinal unit and overall sagittal balance of the cervical spine. *Eur Spine J.* Jan 2008;17(1):20-29.

11. Kim SW, Limson MA, Kim SB, et al. Comparison of radiographic changes after ACDF versus Bryan disc arthroplasty in single and bi-level cases. *Eur Spine J.* Feb 2009;18(2):218-231.

12. Pickett GE, Mitsis DK, Sekhon LH, Sears WR, Duggal N. Effects of a cervical disc prosthesis on segmental and cervical spine alignment. *Neurosurg Focus.* Sep 15 2004;17(3):E5.

13. Fuson RL, Sherman M, Van Vleet J, Wendt T. The conduct of orthopaedic clinical trials. *J Bone Joint Surg Am.* Jul 1997;79(7):1089-1098.

14. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord.* 2006;7:85.

15. Bartels RH, Donk R, Verbeek AL. No justification for cervical disk prostheses in clinical practice: a meta-analysis of randomized controlled trials. *Neurosurgery.* Jun 2010;66(6):1153-1160; discussion 1160.

16. Harrison DE, Harrison DD, Cailliet R, Troyanovich SJ, Janik TJ, Holland B. Cobb method or Harrison posterior tangent method: which to choose for lateral cervical radiographic analysis. *Spine (Phila Pa 1976).* Aug 15 2000;25(16):2072-2078.

17. Toyama Y, Matsumoto M, Chiba K, et al. Realignment of postoperative cervical kyphosis in children by vertebral remodeling. *Spine (Phila Pa 1976).* Nov 15 1994;19(22):2565-2570.

18. Bartels RH, Donk RD, Feuth T. Subsidence of stand-alone cervical carbon fiber cages. *Neurosurgery.* Mar 2006;58(3):502-508; discussion 502-508.

19. Donk RD, Fehlings MG, Verhagen WIM, et al. An assessment of the most reliable method to estimate the sagittal alignment of the cervical spine: analysis of a prospective cohort of 138 cases. *J Neurosurg Spine.* May 2017;26(5):572-576.

20. Anakwenze OA, Auerbach JD, Milby AH, Lonner BS, Balderston RA. Sagittal cervical alignment after cervical disc arthroplasty and anterior cervical discectomy and fusion: results of a prospective, randomized, controlled trial. *Spine (Phila Pa 1976).* Sep 1 2009;34(19):2001-2007.

21. Sasso RC, Metcalf NH, Hipp JA, Wharton ND, Anderson PA. Sagittal alignment after Bryan cervical arthroplasty. *Spine (Phila Pa 1976).* Jun 2011;36(13):991-996.

22. Barrey C, Champain S, Campana S, Ramadan A, Perrin G, Skalli W. Sagittal alignment and kinematics at instrumented and adjacent levels after total disc replacement in the cervical spine. *Eur Spine J.* Aug 2012;21(8):1648-1659.

23. Ahn PG, Kim KN, Moon SW, Kim KS. Changes in cervical range of motion and sagittal alignment in early and late phases after total disc replacement: radiographic follow-up exceeding 2 years. *J Neurosurg Spine.* Dec 2009;11(6):688-695.

24. Di Martino A, Papalia R, Albo E, Cortesi L, Denaro L, Denaro V. Cervical spine alignment in disc arthroplasty: should we change our perspective? *Eur Spine J.* Nov 2015;24 Suppl 7:810-825.

25. Bartels RH, Beems T, Schutte PJ, Verbeek AL. The rationale of postoperative radiographs after cervical anterior discectomy with stand-alone cage for radicular pain. *J Neurosurg Spine.* Mar 2010;12(3):275-279.

26. Iyer S, Nemani VM, Nguyen J, et al. Impact of Cervical Sagittal Alignment Parameters on Neck Disability. *Spine (Phila Pa 1976).* Mar 2016;41(5):371-377.

27. Koerner JD, Kepler CK, Albert TJ. Revision surgery for failed cervical spine reconstruction: review article. *HSS J.* Feb 2015;11(1):2-8.

28. Nemani VM, Derman PB, Kim HJ. Osteotomies in the Cervical Spine. *Asian Spine J.* Feb 2016;10(1):184-195.

Chapter 9

Summary and Considerations

This thesis is a contribution to a long-standing discussion of various methods of anterior cervical discectomy. Since its introduction for the purposes of treating degenerative disc disease in the late 1950s, adaptations of the anterior cervical discectomy with fusion (ACDF) have been further developed. In this study, three methods for anterior cervical discectomy for treating single-level degenerative disk disease causing cervical radiculopathy were compared.

After a brief historical overview in **Chapter 1**, the questions that form the basis of this thesis were posed. For reasons that are not known, ACDF is considered the gold standard treatment, and has been clinically proven a reasonable option and adopted by many spine surgeons. Anterior Cervical Discectomy (ACD) without fusion, however, is still considered a viable alternative for single level cervical degenerative disease. Many alternatives, such as cages, have been developed for the autograft harvested from the iliac crest. The possible disadvantages of fusion include the accelerated degeneration of the disc of the adjacent segment: adjacent segment disease (ASD). In the late 1990s this led to the development and clinical implementation of a new implant—the disc prosthesis or arthroplasty (ACDA). At that time there was confusion over the best surgical treatment for single level degenerative disc disease by anterior approach: ACD, ACDF or ACDA.

In **Chapter 2** we attempt to rate clinical relevance while also discussing the results of a meta-analysis. We analyzed meta-analyses, comparing ACDF and ACDA. Whereas most of them reported statistically significant differences, none of them addressed clinical relevance appropriately. We used minimal clinically important difference (MCID) as a measure. If the study met MCID, the difference would also be clinically relevant. Since none of the studies disclosed any difference as defined by MCID, it was concluded that these differences, although statistically different, were not clinically relevant. Because this study addressed meta-analyses regarding ACDF and ACDA, we felt the aims of our study were supported.

The definition of a good clinical result is addressed in **Chapter 3,** in which the Neck Disability Index (NDI) was used. In individualized medicine it is mandatory that patients be able to interpret data in order to choose a treatment. "A change of NDI" or "an absolute NDI" is generally meaningless to a patient. Therefore, a correlation between the qualification of the clinical situation as rated by the patient and the NDI score was evaluated. The Neck Disability Index (NDI) is a patient self-assessment outcome measurement tool designed to assess disability and is frequently used to evaluate the effects of the treatment of neck-related problems. We correlated the qualification of the situation by patients themselves and NDI. An NDI ffi 7 corresponded to a good outcome according to patients.

The clinical results of the first randomized controlled trial comparing ACD, ACDF and ACDA are presented in **Chapter 4**. Due to the results of a meta-analysis we published, we concluded the trial prematurely. Also, due to the small sample size, statements should be considered inconclusive. However, any difference in clinical outcome was not shown.

In **Chapter 5** we evaluate factors that could possibly be predictive of a good outcome following an ACD procedure as defined in chapter 3. For patients it is very important to be aware of factors that influence treatment outcomes. To counsel patients properly before an operation, providing

information about factors that can contribute to the outcome is important. Predictive factors were determined using NDI, a patient-reported outcome measurement tool. An NDI score ffi 7 was considered a good outcome. Obviously negative predictive factors included smoking and c5c6 as the symptomatic level. Less prominent negative predictive factors were re-operation for adjacent segment disease and, to an even lesser degree, some loss of lordosis. No difference could be established between gender and the anterior techniques employed.

Adjacent segment disease (ASD) is the subject of inquiry in **Chapter 6**. We defined symptomatic ASD as signs and symptoms caused by degeneration of an intervertebral disc adjacent to a level of previous anterior cervical disc surgery. ASD has been predominantly evaluated radiologically. A frequently cited annual rate of ASD is 2.9%. However radiographic evidence of degeneration does not automatically lead to clinical signs or symptoms. In our cohort of 142 patients, we calculated an overall annual rate of symptomatic ASD after ACD of 0.7%. We think the overall annual rate in our study was lower because of our strict preoperative inclusion criteria. We excluded patients with initially obvious radiological findings of disc protrusion or even herniation at an adjacent level. We found no statistically significant correlations between any of the investigated factors and symptomatic ASD except for the surgical method employed. Symptomatic ASD was seen less often in anterior cervical discectomy solely or anterior cervical discectomy with arthroplasty than in anterior cervical discectomy with stand-alone cage. Whether disk prosthesis should be advised so as to prevent adjacent segment disease cannot be concluded based on the results in this study due to the sample size.

An assessment of a method to reliably estimate the sagittal alignment of the cervical spine is presented in **Chapter 7**. A lordosis is the natural shape of the cervical spine and restoration of the lordosis is a goal of most surgeries. The influence of surgery on the sagittal angle of the level of interest but also the angle from C2 to C7 as a measure of global sagittal alignment is currently a subject of debate. In order to develop a reliable method of assessing sagittal alignment of the cervical spine, we reviewed our cervical radiographs that were obtained according to the protocol of the RCT. The Toyama method was modified by drawing a line from the posterior inferior part of the vertebral body of C2 to the posterior upper part of the vertebral body of C7 without any measurements. It proved to be accurate, straightforward, and reliable in qualifying the sagittal alignment of the cervical spine. This technique can be readily and easily incorporated into preoperative surgical planning.

Aspects of cervical sagittal alignment after different anterior discectomy procedures for single level cervical degenerative disc disease were discussed in **Chapter 8**. Although the angle at the involved level became more lordotic after ACDF and more kyphotic after ACD, it tended to normalize to its former preoperative value at approximately week nine of evaluation. For ACD the change was minimal and remained locally kyphotic. At longer follow-up global cervical lordosis was not affected by the procedure for the cervical anterior discectomy. Irrespective of the technique used for anterior cervical discectomy for single level degenerative disc disease, the alignment of the cervical spine remains unaltered.

Work on this thesis afforded the opportunity for further reflections: although we could not definitely conclude which treatment is superior, we are convinced that uniform advice for all patients is not optimal. Prognostic factors should be defined for each method. Since at the time of writing long term follow-up is available to thousands of patients, it is not opportune to design a new trial. A better solution would be to combine available information and assess which patient will benefit from which procedure. This will contribute to furthering patient-tailored health care.

Furthermore, the human body will try to maintain the head in neutral axis in the horizontal plane optimally for the Visio vestibular system and to restore sagittal balance. Restoring local cervical lordosis as an argument for the promotion of a certain procedure for single level cervical degenerative disc disease is at minimum debatable.

DANKWOORD

Dit hoofdstuk zou eigenlijk het grootste moeten zijn, want het is mijns inziens het belangrijkste onderdeel van dit proefschrift. Een ieder die meegewerkt heeft aan de totstandkoming van dit proefschrift zouden hier vermeld moeten worden. Dat is een onmogelijke taak, maar ik ga het toch proberen, wetende dat ik niet volledig kan en zal zijn.
Jullie zijn allemaal trotse eigenaar van dit proefschrift, het is onze gezamenlijke inspanning geweest waar ik mijn naam boven mag zetten .En een inspanning is het geweest en niet alleen voor mij, dat realiseer ik me maar al te goed. Zoals reeds vele promovendi voor mij hebben geschreven is dit niet alleen een proef van bekwaamheid, maar ook een beproeving.

Allemaal een heel hartelijke en welgemeende DANKJEWEL.

Een onderneming waarbij vasthoudendheid op de eerste plaats staat.

Beste Ronald: Je bent een ware (pro)motor, een diesel, eenmaal opgang niet meer te stoppen. Toonbeeld van vasthoudendheid, inspirerende en drijvende kracht, klankbord en stok achter de deur! Altijd aanspreekbaar, corrigerend waar nodig maar nooit gebrek aan ideeën en oplossingen.

Kwaliteit en originaliteit voor jou vanzelfsprekend. Een bezielende (hoog)leraar.

Liefste Ans, als geen ander weet je om te gaan met mijn stress en dat al jaren. Je creëert de noodzakelijke rust en warmte en zorgt voor het juiste perspectief en juiste focus. Heel veel dank voor je onvoorwaardelijke steun.

Lieve Remy en Fabrice, de paranimfen, hoe gaaf is het, met je zoons naast je deze zeer bijzondere dag te beleven! Maxine-Carlijn beleeft deze dag in haar eigen wereldje.

Liefste Els: Bedankt dat ik Ronald mocht "lenen" Je bent met hem al veel gewend, maar naast zijn drukke werkzaamheden heb je hem nog meer moeten afstaan tijdens de begeleiding van mijn proefschrift. Deze is afgerond maar ik vrees en jij weet het, dat de volgende alweer voor de deur staat. De diesel dendert door.

Beste Wim: Letterlijk en figuurlijk betrokken, vanaf het begin. Je frisse kritische blik, input, opmerkingen, aanvullingen, onmisbaar. Een wetenschapper in een drukke perifere praktijk . Chapeau!

Beste Allard: Invalshoek verrassend en origineel. Kort en bondig op een punt gebracht. Je commentaren verbeterend en aanvullend.

Hooggeleerde heer Edwards,

Beste Carmen, beste Barend, beste Jan, voor jullie onmiddellijke bereidheid om plaats te nemen in de manuscript commissie een groot dankjewel. Verheug me erop met jullie van gedachten te wisselen over dit onderwerp.

Dear Dezsö, you didn't hesitate a moment when I asked you to join the manuscript committee, your coming was also not a discussion. Thank you for making the long trip.

Mede auteurs, Hooggeleerde André Verbeek, Weledelgeleerde Hans Groenewoud Dear Professor Michael Fehlings Weledelgeleerde Hisse Arnts, bedankt voor jullie bijdrage, elk specifiek op zijn gebied, onmisbaar.

Beste Marian, Huesken, representatief voor de ok assistenten neurochirurgie en orthopedie, accuraat en betrokken. Definitieve procedure pas kort van te voren bekend, toch slaagde je erin om de ok ´s vlekkeloos te laten verlopen, dankjewel.

Beste Keetie Kremers, Nurse practitioner orthopedie, samen met de poli assistentes zorgde je ervoor dat alle vragen lijsten bij de poli controles werden ingevuld. En ingevuld worden moesten ze van jou. Mijn dank voor je vasthoudendheid.

Secretariaat neurochirurgie, centrale spil, randomiseren, verzamelpunt van alle dossiers, ondersteund waar nodig door secretariaat orthopedie. Dank voor jullie hulp en ondersteuning om dit proefschrift tot een goed einde te brengen.

Vanaf het ontstaan van het idee, de uitwerking, de uiteindelijke start van het onderzoek en uiteindelijk de analyse van de gegevens heeft u allen meer dan een steentje bijgedragen.

De spontaniteit en de inzet was voor velen zelfs vanzelfsprekend.

Het moge duidelijk zijn dat dit niet vanzelfsprekend is, het heeft ook veel vrije tijd gekost.

Bijzondere dank gaat uit naar alle patiënten die letterlijk hun nek hebben uitgestoken en opgenomen zijn in dit onderzoek.

Uiteindelijk zorgt het secretariaat van ViaSana ervoor dat de laatste puntjes op de i gezet kunnen worden en het proefschrift op tijd kan worden afgeleverd.

Lieve Wendy en Wim, beiden een groot hart op de juiste plaats. Liever op de achtergrond, maar jullie bijdrage groot en klein, verdient een groot dankjewel. Dit boekje werd door jullie mogelijk gemaakt.

Er is een aanzienlijke kans dat ik iemand vergeet te bedanken. Iedereen die in welke opzicht dan ook heeft geholpen dank ik hiervoor.

Curriculum Vitae

Roland Donk werd geboren op 2 april 1956 in de gemeente Venlo. Als jongste in een gezin van vijf groeide hij op met 2 oudere zussen en 2 oudere broers in het stadsdeel Blerick van de gemeente Venlo. Hier bezocht hij ook het Blaricum College en sloot deze in 1974 af.

Direct aansluitend volgde een onvergetelijk tijd aan de toen nog Katholieke Universiteit Nijmegen. De studie geneeskunde werd in 1981 afgesloten en de opleiding in Duitsland begon. De basis werd gelegd in Duisburg, algemene Heelkunde bij Dr. K. Partenheimer.

Een opleiding plaats bij Prof. Dr. K-F. Schlegel aan het Uniklinikum Essen werd verkregen. In voorbereiding daarop werd eerst een vooropleiding verlangt in de biomechanica bij Prof. Dr. B. Kummer, in Keulen. Wegens capaciteit gebrek werd een alternatief gezocht en gevonden in de wervelkolomchirurgie bij Dr. Klaus Zielke in het Deutsche Skoliose Centrum in Bad Wildungen. Hier werd de wervelkolomchirurgie met beide armen omhelst en nooit meer losgelaten, een geweldige periode. Na het afronden van de opleiding tot orthopedisch chirurg startte hij samen met Klaus Pickhardt een praktijk in Hattingen Duitsland. De liefde voor de wervelkolomchirurgie deed hem terug keren naar Nederland en werd staflid orthopedie in het Canisius Wilhelmina Ziekenhuis te Nijmegen. De wervelkolomchirurgie werd hier gezamenlijk met de neurochirurgen uitgeoefend en verder ontwikkeld, een intensieve samenwerking. De landelijk toename van deze samenwerking leidde tot de oprichting van de Dutch Spine society. In 2002 werden de werkgroepen spinale chirurgie van de Nederlandse Vereniging van Neurochirurgen en de wervelkolom werkgroep van de Nederlandse Orthopedische Vereniging samengevoegd tot de Dutch Spine Society, waarvan hij de eerste voorzitter werd.

In 2012 verruilde hij het Canisius Wilhelmina Ziekenhuis voor Kliniek ViaSana te Mill.

Ans de Brouwer werd in 1989 zijn echtgenote en samen hebben ze 3 kinderen, Remy, Maxine-Carlijn en Fabrice.

Hun thuis is en was altijd Venlo, "stedje van lol en plezeer" De voltooiing van dit proefschrift moet U dan ook zien in het licht van de wapenspreuk van Venlo "Festina lente, cauta fac omnia mente", vrij vertaald: "Haast je langzaam, doe alles na rijp overleg".